

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2002, 10:19:23 ; Search time 35 Seconds
(without alignments)
87.565 Million cell updates/sec

Title: US-09-812-502A-16
Perfect score: 120
Sequence: 1 XCPXXEKKNDRICTNCCAGXKG 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	115	95.8	368	15	AA854135
2	78	65.0	153	13	AA20099
3	74	61.7	13	15	AA854131
4	74	61.7	58	15	AA854125
5	74	61.7	58	15	AA854126
6	73	60.8	13	15	AA854132
7	73	60.8	54	15	AA854130
8	73	60.8	58	15	AA854127
9	73	60.8	58	15	AA854128
10	73	60.8	58	15	AA854129

11	53	44.2	41	21	AA856552	Arabidopsis thalia
12	53	44.2	51	21	AA856551	Arabidopsis thalia
13	53	44.2	66	21	AA856550	Arabidopsis thalia
14	53	44.2	417	22	AB860155	Drosophila melanog
15	52	43.3	609	21	AA851616	Arabidopsis thalia
16	52	43.3	710	21	AA851615	Arabidopsis thalia
17	52	43.3	721	21	AA851614	Arabidopsis thalia
18	49	40.8	113	21	AA854118	Pinus radiata cell
19	49	40.8	380	22	AB809344	Novel human diago
20	49	40.8	919	23	ABP35687	Fungal ZBC protein
21	49	40.8	3443	20	AAW84559	Human angiotensin
22	47	39.2	171	23	AB895595	Human angiotensin
23	47	39.2	171	23	AB894989	Human angiotensin
24	47	39.2	176	19	AAW80254	Human angiotensin
25	47	39.2	231	23	ABJ01955	Human angiotensin
26	47	39.2	327	22	ABG14985	Human angiotensin
27	47	39.2	904	22	ABG06694	Human angiotensin
28	47	39.2	904	22	ABG15681	Human angiotensin
29	47	39.2	904	22	ABG16968	Human angiotensin
30	46	38.3	103	22	ABG02782	Human angiotensin
31	46	38.3	187	22	AAW89056	Human angiotensin
32	46	38.3	255	22	AAU43158	Human angiotensin
33	46	38.3	318	22	ABG30364	Human angiotensin
34	46	38.3	1329	23	AAU91279	Human angiotensin
35	45	37.5	78	22	AAU42830	Human angiotensin
36	44.5	37.1	1576	21	AA819802	Human laminin 2 ma
37	44.5	37.1	1576	21	AA848453	Human laminin 8 po
38	44.5	37.1	1576	23	AB881595	Human laminin 10 t
39	44.5	37.1	1584	21	AA819804	Human laminin 2 ga
40	44.5	37.1	1609	19	AAW50898	Human laminin 2 ga
41	44.5	37.1	1609	21	AA819801	Human laminin 2 ga
42	44.5	37.1	1609	21	AA848452	Human laminin 8 po
43	44.5	37.1	1609	23	AB881594	Human laminin 10 t
44	44.5	37.1	1617	21	AA819803	Human laminin 2 ga
45	44	36.7	258	22	AA850082	Rat TNFR (p80) ext

ALIGNMENTS

RESULT 1
AA854135
ID AA854135 standard; Protein; 368 AA.
AC AA854135;
XX
DT 02-MAR-1995 (first entry)
XX
DE Deduced sequence of mature N-alata PI precursor.
XX
KW Type II serine proteinase, inhibitor precursor; PI; tobacco;
XX transgenic plant; anti-pathogen; anti-predator.
XX Nicotiana alata.

Key	Location/Qualifiers
FT Active-site	5..6
FT Active-site	/label= reactive site
FT Active-site	63..64
FT Active-site	/label= reactive site
FT Active-site	121..122
FT Active-site	/label= reactive site
FT Active-site	179..180
FT Active-site	/label= reactive site
FT Active-site	237..238
FT Active-site	/label= reactive site
FT Active-site	295..296
FT Active-site	/label= reactive site
FT Domain	1..58
FT Domain	/label= 1
FT Domain	59..116
FT Domain	/label= 2
FT Domain	117..174

FT Domain /label= 3
FT 175..232 /label= 4
FT Domain 233..290 /label= 5
FT Domain 291..343 /label= 6
FT Peptide 1..24 /label= 1
FT /label= 2
FT Peptide 25..82 /label= 3
FT Peptide 83..140 /label= 4
FT Peptide 141..198 /label= 5
FT Peptide 199..256 /label= 6
FT Peptide 257..314 /label= 7
FT Peptide 315..368 /label= 7

PN MO9413810-A.
XX 23-JUN-1994.
XX
PF 16-DEC-1993; 93WO-AU00659.
XX
PR 16-DEC-1992; 92AU-0006399.
XX
PA (UWME) UNIV MELBOURNE.
PI
PI Anderson MA, Atkinson AH, Clarke AE, Heath RL,
XX WPI, 1994-21786/26.
DR N-PSDB; AAQ68728, AAQ68729.
XX
XX Nicotiana alata type II serine protease inhibitor precursor and
PT DNA - useful in prodn of anti-pathogen or anti-predator
PT constructs for plants.

PS Claim 16; Page 45-47; 83pp; English.

XX A cDNA library, prep'd. from mRNA from the stigmas and styles of
XX mature flowers of N. alata was screened for clones of highly
XX expressed genes which were not associated with self-incompatibility
XX genotype. Clones encoding a protein with some identity to the type
XX I1 proteinase inhibitors from potato and tomato were selected. The
XX largest clone, NA-P1-2, is given in AAQ68729. The predicted AA sequence
XX in AAQ54135. AAQ68728 is the coding region of AAQ68729. The type II
XX serine PI has six conserved domains wherein the first two are 100%
XX identical and contain chymotrypsin-specific sites. The 3rd, 4th and
XX 5th domains share 95-98% identity and have sites specific for
XX trypsin. The 6th domain also has a trypsin specific site but less
XX identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
XX a divergent 3' sequence. It has a mol. wt. of approx. 42-45kDa
XX with an approx. 29 AA signal sequence. The N-terminal sequence
XX of the monomeric PI is represented in each of the six repeated
XX domains in the predicted sequence of the PI precursor protein.
XX Thus, it is likely that the PI precursor protein is cleaved at six
XX sites to produce seven peptides. Six of the seven peptides,
XX peptides 2-7 (AAQ54125-R54130) would be in the same mol. wt. range
XX as the monomeric PI (about 6kDa) and would have the same N-terminal
XX sequence. Peptide 7 does not contain a consensus site for trypsin
XX or chymotrypsin. Peptide 1 (AAQ54124) is small than 6kDa, has a
XX different N-terminus and was not detected in a purified monomeric
XX PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
XX form a functional proenzyme inhibitor with the inhibitory site on
XX peptide 1 held in the correct conformation by disulfide bonds
XX between the two peptides.

SQ Sequence 368 AA;

```

Query Match          95.8%; Score 115; DB 15; Length 368;
Best Local Similarity 86.4%; Pred: No. 8-08;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      2 CPXXEKKNDRICTNCCAGXKG 23
        ||| ||||| ||||| |||
Db       16 CPRSEKKNDRICTNCCAGTGG 37

RESULT 2
AAR20099
AAC AAR20099 standard; Protein; 153 AA.
AC AAR20099;
AD
AE
AF 24-MAR-1992 (first entry)
AG
AH
AI Wound-inducible potato protease inhibitor IIR.
AJ methyl jasmonate; transgenic plant; predator defence protein.
AK Solanum tuberosum.
AL
AM Key Location/Qualifiers
AN Peptide 1..30
AO FT /label= transit
AP FT 31..153
AQ FT /label= Inhibitor_IIR
AR WO9118512-A.
AS
AT PN 12-DEC-1991.
AU PD 24-MAY-1991; 91WO-US03685.
AV PF 25-MAY-1990; 90US-0528956.
AW XX
AX PR (UNITW ) WASHINGTON STATE UN.
AY PA Ryan CA, Farmer EE;
AZ PI WPI: 1992-007120/01.
BA DR N-PsDB; AAQ20253.
BB
BC Induction of defence protein prodn. in plants - by treatment with
BD agent capable of such induction, esp. usmasonic acid cpd.
BE XX
BF PS Example 7; Fig 6; 40pp; English.
BG XX
BH A Russert Burbank potato genomic library was screened by using nick-
BI translated wound-induced tomato inhibitor II cDNA as a probe.
BJ CC Positive clones were rescreened and a clone contg. an 8bp EcoRI
BK insert hybridised most strongly with the probe. A TaqI 2.6kbp
BL fragment of this insert was subcloned in M13 and a 1.24kbp fragment
BM sequenced. The amino acid sequence of the inhibitor was deduced
BN from the nucleotide sequence. The nascent protein is assumed to be
BO processed during or after synthesis between amino acid residues 30
BP and 31 to produce the mature inhibitor.
BQ CC
BR XX
BS SQ Sequence 153 AA;

Query Match          65.0%; Score 78; DB 13; Length 153;
Best Local Similarity 59.1%; Pred. No. 0.0039;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY      2 CPXXEKKNDRICTNCCAGXKG 23
        ||| | : ||||| |||
Db       45 CPRSEGSPENDICTNCCAGYKG 66

RESULT 3
AAR54131
AA AAR54131 standard; Peptide; 13 AA.
```

```

XX AARS4131;
XX AC
XX DT 02-MAR-1995 (first entry)
XX DE
XX DE N-terminal sequence of 6kD PI protein.
XX KW
XX KW Type II serine proteinase inhibitor precursor; PI; tobacco;
XX KW transgenic plant; anti-pathogen; anti-predator; peptide.
XX OS Nicotiana alata.
XX PN WO9413810-A.
XX PD
XX PD 23-JUN-1994.
XX PF 16-DEC-1993; 93WO-AU00659.
XX PR 16-DEC-1992; 92AU-0006399.
XX PA (UYME ) UNIV MELBOURNE.
XX PI Anderson MA, Atkinson AH, Clarke AE, Heath RL;
XX WPI; 1994-217886/26.
XX Nicotiana alata type II serine protease inhibitor precursor and
XX DNA - useful in prodn of anti-pathogen or anti-predator
XX constructs for plants.
XX PS Example; Page 28; 83pp; English.
XX Stigmas of N. alata were extracted. A protein with inhibitory
XX activity of both trypsin and chymotrypsin was eluted. The PI
XX activity co-eluted with a protein of about 6kD. The N-terminal
XX sequence (AARS4131/R54132) was obtd. from the purified PI protein.
XX This sequence of AAs corresp. to six regions in the deduced
XX sequence (AARS4135) of the cDNA clone (AA068729), starting as poens
XX 25,83,141,199,257 and 315. At poen. 11 of the N-terminal sequence
XX both T and K were detected. This is consistent with the purified
XX inhibitor comprising a mixture of six peptides (see AARS4135).
XX SQ Sequence 13 AA;
Query Match 61.7%; Score 74; DB 15; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 11 DRICNTCCAGXKG 23
Db 1 DRICNTCCAGTKG 13
RESULT 4
AARS4125
ID AARS4125 standard; Peptide; 58 AA.
XX AARS4125;
XX AC
XX DT 02-MAR-1995 (first entry)
XX DE
XX DE Sequence of peptide 2 of N-alata PI precursor.
XX KW
XX KW Type II serine proteinase inhibitor precursor; PI; tobacco;
XX KW transgenic plant; anti-pathogen; anti-predator; peptide.
XX OS Nicotiana alata.
XX PN WO9413810-A.
XX PD
XX PD 23-JUN-1994.
XX PF 16-DEC-1993; 93WO-AU00659.

```

```

XX 16-DEC-1992; 92AU-0006399.
XX (UYME ) UNIV MELBOURNE.
XX Anderson MA, Atkinson AH, Clarke AE, Heath RL;
XX WPI; 1994-217886/26.
XX N-PSDB; AAQ68728, AAQ68729.
XX Nicotiana alata type II serine protease inhibitor precursor and
XX DNA - useful in prodn of anti-pathogen or anti-predator
XX constructs for plants.
XX Claim 18; Page 50; 83pp; English.
XX A cDNA library, prep'd. from mRNA from the stigmas and styles of
XX mature flowers of N. alata was screened for clones of highly
XX expressed genes which were not associated with self-incompatibility
XX genotype. Clones encoding a protein with some identity to the type
XX II proteinase inhibitors from potato and tomato were selected. The
XX largest clone, NA-EI-2, is given in AAQ68729. The predicted AA sequence
XX in AARS4135. AAQ68728 is the coding region of AAQ68729. The type II
XX serine PI has six conserved domains wherein the first two are 100%
XX identical and contain chymotrypsin-specific sites. The 3rd, 4th and
XX 5th domains share 95-98% identity and have sites specific for
XX trypsin. The 6th domain also has a trypsin specific site but less
XX identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
XX a divergent 3' sequence. It has a mol. wt. of approx. 42-45kDa
XX of the monomeric PI is represented in each of the six repeated
XX domains in the predicted sequence of the PI precursor protein.
XX Thus, it is likely that the PI precursor protein is cleaved at six
XX sites to produce seven peptides. Six of the seven peptides,
XX peptides 2-7 (AARS4125-R54130) would be in the same mol. wt. range
XX as the monomeric PI (about 6kDa) and would have the same N-terminal
XX or chymotrypsin. Peptide 1 (AARS4124) is small than 6kDa, has a
XX different N-terminus and was not detected in a purified monomeric
XX PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
XX form a functional proenzyme inhibitor with the inhibitory site on
XX peptide 1 held in the correct conformation by disulphide bonds
XX between the two peptides. A monomer which corresp. to any of the
XX peptides 2-7 is claimed.
XX SQ Sequence 58 AA;
Query Match 61.7%; Score 74; DB 15; Length 58;
Best Local Similarity 92.3%; Pred. No. 0.0057;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 11 DRICNTCCAGXKG 23
Db 1 DRICNTCCAGTKG 13
RESULT 5
AARS4126
ID AARS4126 standard; Peptide; 58 AA.
XX AARS4126;
XX AC
XX DT 02-MAR-1995 (first entry)
XX DE
XX DE Sequence of peptide 3 of N-alata PI precursor.
XX KW
XX KW Type II serine proteinase inhibitor precursor; PI; tobacco;
XX KW transgenic plant; anti-pathogen; anti-predator; peptide.
XX OS Nicotiana alata.
XX PN WO9413810-A.
XX PD
XX PD 16-DEC-1993; 93WO-AU00659.

```

PD 23-JUN-1994.
 XX PF 16-DEC-1993; 93WO-AU00659.
 XX PR 16-DEC-1992; 92AU-0006399.
 XX (UYME) UNIV MELBOURNE.
 XX PA Anderson MA, Atkinson AH, Clarke AE, Heath RL,
 XX PI WPI; 1994-217886/26.
 XX DR N-PSDB; AAQ68728, AAQ68729.
 XX PT Nicotiana glauca type II serine protease inhibitor precursor and
 PT DNA - useful in prodn of anti-pathogen or anti-predator
 PT constructs for plants.
 PS Claim 18; Page 51; 83pp; English.

CC A cDNA library, prep'd. from mRNA from the stigmas and styles of
 CC mature flowers of N. glauca was screened for clones of highly
 CC expressed genes which were not associated with self-incompatibility
 CC genotype. Clones encoding a protein with some identity to the type
 CC I1 proteinase inhibitors from potato and tomato were selected. The
 CC largest clone, NA-PI-2, is given in AAQ68729. The predicted AA sequence
 CC in AAQ68728 is the coding region of AAQ68729. The type II
 CC serine PI has six conserved domains wherein the first two are 100%
 CC identical and contain chymotrypsin-specific sites. The 3rd, 4th and
 CC 5th domains share 95-98% identity and have sites specific for
 CC trypsin. The 6th domain also has a trypsin specific site but less
 CC identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
 CC a divergent 3' sequence. It has a mol. wt. of approx. 42-45kDa
 CC with an approx. 29 AA signal sequence. The N-terminal sequence
 CC of the monomeric PI is represented in each of the six repeated
 CC domains in the predicted sequence of the PI precursor protein.
 CC Thus, it is likely that the PI precursor protein is cleaved at six
 CC sites to produce seven peptides. Six of the seven peptides, w/
 CC peptides 2-7 (AAQ68728-854130) would be in the same mol. wt. range
 CC as the monomeric PI (about 6kDa) and would have the same N-terminal
 CC or chymotrypsin. Peptide 1 (AAQ68724) is small than 6kDa, has a
 CC different N-terminus and was not detected in a purified monomeric
 CC PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
 CC form a functional proteinase inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds
 CC between the two peptides. A monomer which corresp. to any of the
 CC peptides 2-7 is claimed.

SO Sequence 58 AA;
 Query Match 61.7%; Score 74; DB 15; Length 58;
 Best Local Similarity 92.3%; Pred. No. 0.0057;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 DRICNCCAGKKG 23
 DB 1 DRICNCCAGKKG 13

RESULT 6
 ID AAR54132 standard; Peptide; 13 AA.
 XX AAR54132;
 AC 02-MAR-1995 (first entry)
 DT N-terminal sequence of 6KD PI protein.
 DE Type II serine proteinase inhibitor precursor; PI; tobacco;
 XX transgenic plant; anti-pathogen; anti-predator; peptide.
 KW Nicotiana glauca.
 XX OS

XX PN W09413810-A.
 XX PD 23-JUN-1994.
 XX PF 16-DEC-1993; 93WO-AU00659.
 XX PR 16-DEC-1992; 92AU-0006399.
 XX (UYME) UNIV MELBOURNE.
 XX PA Anderson MA, Atkinson AH, Clarke AE, Heath RL,
 XX PI WPI; 1994-217886/26.
 XX DR Nicotiana glauca type II serine protease inhibitor precursor and
 XX PT DNA - useful in prodn of anti-pathogen or anti-predator
 XX PT constructs for plants.
 XX PS Example; Page 28; 83pp; English.

CC Stigmas of N. glauca were extracted. A protein with inhibitory
 CC activity of both trypsin and chymotrypsin was eluted. The PI
 CC activity co-eluted with a protein of about 6kD. The N-terminal
 CC sequence (AAR54131/R54132) was obtd. from the purified PI protein.
 CC This sequence of AAs corresp. to six regions in the deduced
 CC sequence (AAR54135) of the cDNA clone (AAQ68729), starting as posns
 CC 25, 83, 141, 199, 257 and 315. At posn. 11 of the N-terminal sequence
 CC both T and K were detected. This is consistent with the purified
 CC inhibitor comprising a mixture of six peptides (see AAR54135).

SO Sequence 13 AA;
 Query Match 60.8%; Score 73; DB 15; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.0021;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 DRICNCCAGKKG 23
 DB 1 DRICNCCAGKKG 13

RESULT 7
 ID AAR54130 standard; Peptide; 54 AA.
 XX AAR54130;
 AC 02-MAR-1995 (first entry)
 DT Sequence of peptide 7 of N-glauca PI precursor.
 DE Type II serine proteinase inhibitor precursor; PI; tobacco;
 XX transgenic plant; anti-pathogen; anti-predator; peptide.
 KW Nicotiana glauca.
 XX OS
 XX PN W09413810-A.
 XX PD 23-JUN-1994.
 XX PF 16-DEC-1993; 93WO-AU00659.
 XX PR 16-DEC-1992; 92AU-0006399.
 XX (UYME) UNIV MELBOURNE.
 XX PA Anderson MA, Atkinson AH, Clarke AE, Heath RL,
 XX PI WPI; 1994-217886/26.
 XX DR N-PSDB; AAQ68728, AAQ68729.
 XX PT Nicotiana glauca type II serine protease inhibitor precursor and

PT DNA - useful in prodn of anti-pathogen or anti-predator
 PT constructs for plants.
 XX
 PS Claim 18; Page 53; 83pp; English.
 XX
 CC A cDNA library, prep'd. from mRNA from the stigmas and styles of
 CC mature flowers of N. alata was screened for clones of highly
 CC expressed genes which were not associated with self-incompatibility
 CC genotype. Clones encoding a protein with some identity to the type
 CC II proteinase inhibitors from potato and tomato were selected. The
 CC largest clone, NA-PI-2, is given in AAQ68729. The predicted AA sequence
 CC in AAR54135. AAQ68728 is the coding region of AAQ68729. The type II
 CC serine PI has six conserved domains wherein the first two are 100%
 CC identical and contain chymotrypsin-specific sites. The 3rd, 4th and
 CC 5th domains share 95-98% identity and have sites specific for
 CC trypsin. The 6th domain also has a trypsin specific site but less
 CC identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
 CC a divergent 3' sequence. It has a mol. st. of approx. 42-45kDa
 CC of the monomeric PI is represented in each of the six repeated
 CC domains in the predicted sequence of the PI precursor protein.
 CC Thus, it is likely that the PI precursor protein is cleaved at six
 CC sites to produce seven peptides. Six of the seven peptides,
 CC peptides 2-7 (AAR54125-R54130) would be in the same mol. wt. range
 CC as the monomeric PI (about 6kDa) and would have the same N-terminal
 CC sequence. Peptide 7 does not contain a consensus site for trypsin
 CC or chymotrypsin. Peptide 1 (AAR54124) is small than 6kDa, has a
 CC different N-terminus and was not detected in a purified monomeric
 CC PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
 CC form a functional proteinase inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds
 CC between the two peptides. A monomer which corresp. to any of the
 CC peptides 2-7 is claimed.
 XX
 XX Sequence 54 AA;
 SQ
 Query Match 60.8%; Score 73; DB 15; Length 54;
 Best Local Similarity 92.3%; Pred. No. 0.0074;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY, 11 DRICNCCAGKKG 23
 Db 1 DRICNCCAGKKG 13
 RESULT 8
 AAR54127
 ID AAR54127 standard; Peptide; 58 AA.
 AC AAR54127;
 XX
 DT 02-MAR-1995 (first entry)
 DE Sequence of peptide 4 of N-alata PI precursor.
 XX
 KW Type II serine proteinase inhibitor precursor; PI; tobacco;
 KW transgenic plant; anti-pathogen; anti-predator; peptide.
 XX
 OS Nicotiana alata.
 XX
 PN WO9413810-A.
 XX
 PD 23-JUN-1994.
 XX
 PF 16-DEC-1993; 93WO-AU00659.
 XX
 PR 16-DEC-1992; 92AU-0006399.
 XX
 PA (UYME) UNIV MELBOURNE.
 XX
 PI Anderson MA, Atkinson AH, Clarke AE, Heath RL;
 XX WPI; 1994-217886/26.

DR N-PSDB; AAQ68728, AAQ68729.
 XX
 PT Nicotiana alata type II serine protease inhibitor precursor and
 PT DNA - useful in prodn of anti-pathogen or anti-predator
 PT constructs for plants.
 XX
 PS Claim 18; Page 51; 83pp; English.
 XX
 CC A cDNA library, prep'd. from mRNA from the stigmas and styles of
 CC mature flowers of N. alata was screened for clones of highly
 CC expressed genes which were not associated with self-incompatibility
 CC genotype. Clones encoding a protein with some identity to the type
 CC II proteinase inhibitors from potato and tomato were selected. The
 CC largest clone, NA-PI-2, is given in AAQ68729. The predicted AA sequence
 CC in AAR54135. AAQ68728 is the coding region of AAQ68729. The type II
 CC serine PI has six conserved domains wherein the first two are 100%
 CC identical and contain chymotrypsin-specific sites. The 3rd, 4th and
 CC 5th domains share 95-98% identity and have sites specific for
 CC trypsin. The 6th domain also has a trypsin specific site but less
 CC identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
 CC a divergent 3' sequence. It has a mol. st. of approx. 42-45kDa
 CC of the monomeric PI is represented in each of the six repeated
 CC domains in the predicted sequence of the PI precursor protein.
 CC Thus, it is likely that the PI precursor protein is cleaved at six
 CC sites to produce seven peptides. Six of the seven peptides,
 CC peptides 2-7 (AAR54125-R54130) would be in the same mol. wt. range
 CC as the monomeric PI (about 6kDa) and would have the same N-terminal
 CC sequence. Peptide 7 does not contain a consensus site for trypsin
 CC or chymotrypsin. Peptide 1 (AAR54124) is small than 6kDa, has a
 CC different N-terminus and was not detected in a purified monomeric
 CC PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
 CC form a functional proteinase inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds
 CC between the two peptides. A monomer which corresp. to any of the
 CC peptides 2-7 is claimed.
 XX
 XX Sequence 58 AA;
 SQ
 Query Match 60.8%; Score 73; DB 15; Length 58;
 Best Local Similarity 92.3%; Pred. No. 0.0078;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY, 11 DRICNCCAGKKG 23
 Db 1 DRICNCCAGKKG 13
 RESULT 9
 AAR54128
 ID AAR54128 standard; Peptide; 58 AA.
 AC AAR54128;
 XX
 DT 02-MAR-1995 (first entry)
 DE Sequence of peptide 5 of N-alata PI precursor.
 XX
 KW Type II serine proteinase inhibitor precursor; PI; tobacco;
 KW transgenic plant; anti-pathogen; anti-predator; peptide.
 XX
 OS Nicotiana alata.
 XX
 PN WO9413810-A.
 XX
 PD 23-JUN-1994.
 XX
 PF 16-DEC-1993; 93WO-AU00659.
 XX
 PR 16-DEC-1992; 92AU-0006399.
 XX
 PA (UYME) UNIV MELBOURNE.

PI Anderson MA, Atkinson AH, Clarke AE, Heath RL;
 XX WPI: 1994-217886/26.
 XX DR N-PSDB; AAQ68728, AAQ68729.
 XX Nicotiana glauca type II serine protease inhibitor precursor and
 PT DNA - useful in prodn of anti-pathogen or anti-predator
 PT constructs for plants.
 PS Claim 18; Page 52; 83pp; English.
 CC A cDNA library, prep'd. from mRNA from the stigmas and styles of
 CC mature flowers of N. glauca was screened for clones of highly
 CC expressed genes which were not associated with self-incompatibility
 CC genotype. Clones encoding a protein with some identity to the type
 CC II proteinase inhibitors from potato and tomato were selected. The
 CC largest clone, NA-PI-2, is given in AAQ68729. The predicted AA sequence
 CC in AAQ68728 is the coding region of AAQ68729. The type II
 CC serine PI has six conserved domains wherein the first two are 100%
 CC identical and contain chymotrypsin-specific sites. The 3rd, 4th and
 CC 5th domains share 95-98% identity and have sites specific for
 CC trypsin. The 6th domain also has a trypsin specific site but less
 CC identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
 CC a divergent 3' sequence. It has a mol. wt. of approx. 42-45kDa
 CC with an approx. 29 AA signal sequence. The N-terminal sequence
 CC of the monomeric PI is represented in each of the six repeated
 CC domains in the predicted sequence of the PI precursor protein.
 CC Thus, it is likely that the PI precursor protein is cleaved at six
 CC sites to produce seven peptides. Six of the seven peptides,
 CC peptides 2-7 (AAQ68728-54130) would be in the same mol. wt. range
 CC as the monomeric PI (about 6kDa) and would have the same N-terminal
 CC sequence. Peptide 7 does not contain a consensus site for trypsin
 CC or chymotrypsin. Peptide 1 (AAQ68728) is small than 6kDa, has a
 CC different N-terminus and was not detected in a purified monomeric
 CC PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
 CC form a functional proteinase inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds
 CC between the two peptides. A monomer which corresp. to any of the
 CC peptides 2-7 is claimed.
 CC XX
 SQ Sequence 58 AA;
 Query Match 60.8%; Score 73; DB 15; Length 58;
 Best Local Similarity 92.3%; Pred. No. 0.0078;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 11 DRICNCCAGKKG 23
 DB 1 DRICNCCAGKKG 13
 RESULT 10
 AAR54129
 ID AAR54129 standard; Peptide; 58 AA.
 AC AAR54129;
 DT 02-MAR-1995 (first entry)
 XX Sequence of peptide 6 of N-alata PI precursor.
 DE Type II serine proteinase inhibitor precursor; PI, tobacco;
 KW transgenic plant; anti-pathogen; anti-predator; peptide.
 OS Nicotiana glauca.
 XX WO9413810-A.
 PN 23-JUN-1994.
 XX 16-DEC-1993; 93WO-AU00659.
 PF 16-DEC-1992; 92AU-0006399.
 XX PR

XX (UYME) UNIV MELBOURNE.
 PA Anderson MA, Atkinson AH, Clarke AE, Heath RL;
 XX WPI: 1994-217886/26.
 XX DR N-PSDB; AAQ68728, AAQ68729.
 XX Nicotiana glauca type II serine protease inhibitor precursor and
 PT DNA - useful in prodn of anti-pathogen or anti-predator
 PT constructs for plants.
 PS Claim 18; Page 52; 83pp; English.
 CC A cDNA library, prep'd. from mRNA from the stigmas and styles of
 CC mature flowers of N. glauca was screened for clones of highly
 CC expressed genes which were not associated with self-incompatibility
 CC genotype. Clones encoding a protein with some identity to the type
 CC II proteinase inhibitors from potato and tomato were selected. The
 CC largest clone, NA-PI-2, is given in AAQ68729. The predicted AA sequence
 CC in AAQ68728 is the coding region of AAQ68729. The type II
 CC serine PI has six conserved domains wherein the first two are 100%
 CC identical and contain chymotrypsin-specific sites. The 3rd, 4th and
 CC 5th domains share 95-98% identity and have sites specific for
 CC trypsin. The 6th domain also has a trypsin specific site but less
 CC identity to the 3rd, 4th and 5th domains (79-90%) due mainly to
 CC a divergent 3' sequence. It has a mol. wt. of approx. 42-45kDa
 CC with an approx. 29 AA signal sequence. The N-terminal sequence
 CC of the monomeric PI is represented in each of the six repeated
 CC domains in the predicted sequence of the PI precursor protein.
 CC Thus, it is likely that the PI precursor protein is cleaved at six
 CC sites to produce seven peptides. Six of the seven peptides,
 CC peptides 2-7 (AAQ68728-54130) would be in the same mol. wt. range
 CC as the monomeric PI (about 6kDa) and would have the same N-terminal
 CC sequence. Peptide 7 does not contain a consensus site for trypsin
 CC or chymotrypsin. Peptide 1 (AAQ68728) is small than 6kDa, has a
 CC different N-terminus and was not detected in a purified monomeric
 CC PI prep'n. It could be envisaged that peptide 1 and peptide 7 would
 CC form a functional proteinase inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds
 CC between the two peptides. A monomer which corresp. to any of the
 CC peptides 2-7 is claimed.
 CC XX
 SQ Sequence 58 AA;
 Query Match 60.8%; Score 73; DB 15; Length 58;
 Best Local Similarity 92.3%; Pred. No. 0.0078;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 11 DRICNCCAGKKG 23
 DB 1 DRICNCCAGKKG 13
 RESULT 11
 AAG56552
 ID AAG56552 standard; Protein; 41 AA.
 AC AAG56552;
 DT 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 72713.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS EP1033405-A2.
 PN 06-SEP-2000.
 XX PD

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.2%; Score 53; DB 21; Length 41;
Best Local Similarity 47.4%; Pred. No. 2.8;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCCAG 20
6 CSKKKKKKSSCCCCCKG 24
Db

RESULT 12
ID AAG56551 standard; Protein; 51 AA.
AAG56551

XX AAG56551;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 72712.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.


```
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.

PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161392.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.2%; Score 53; DB 21; Length 51;
Best Local Similarity 47.4%; Pred. No. 3.4;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2 CPXXEKKNDRICTNCCAG 20
Db 16 CSKKKKKKDSCCCCKG 34

RESULT 13
AAG56550
ID AAG56550 standard; Protein; 66 AA.
XX
AC AAG56550;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 72711.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
```

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139839.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140625.
PR 24-JUN-1999; 99US-0140623.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150584.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155113.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158065.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 44.2%; Score 53; DB 21; Length 66;
 Best Local Similarity 47.4%; Pred. No. 4.3;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXEEKNDRICTNCCAG 20
 Db 31 CSKKKKKKSSCCCCCKG 49

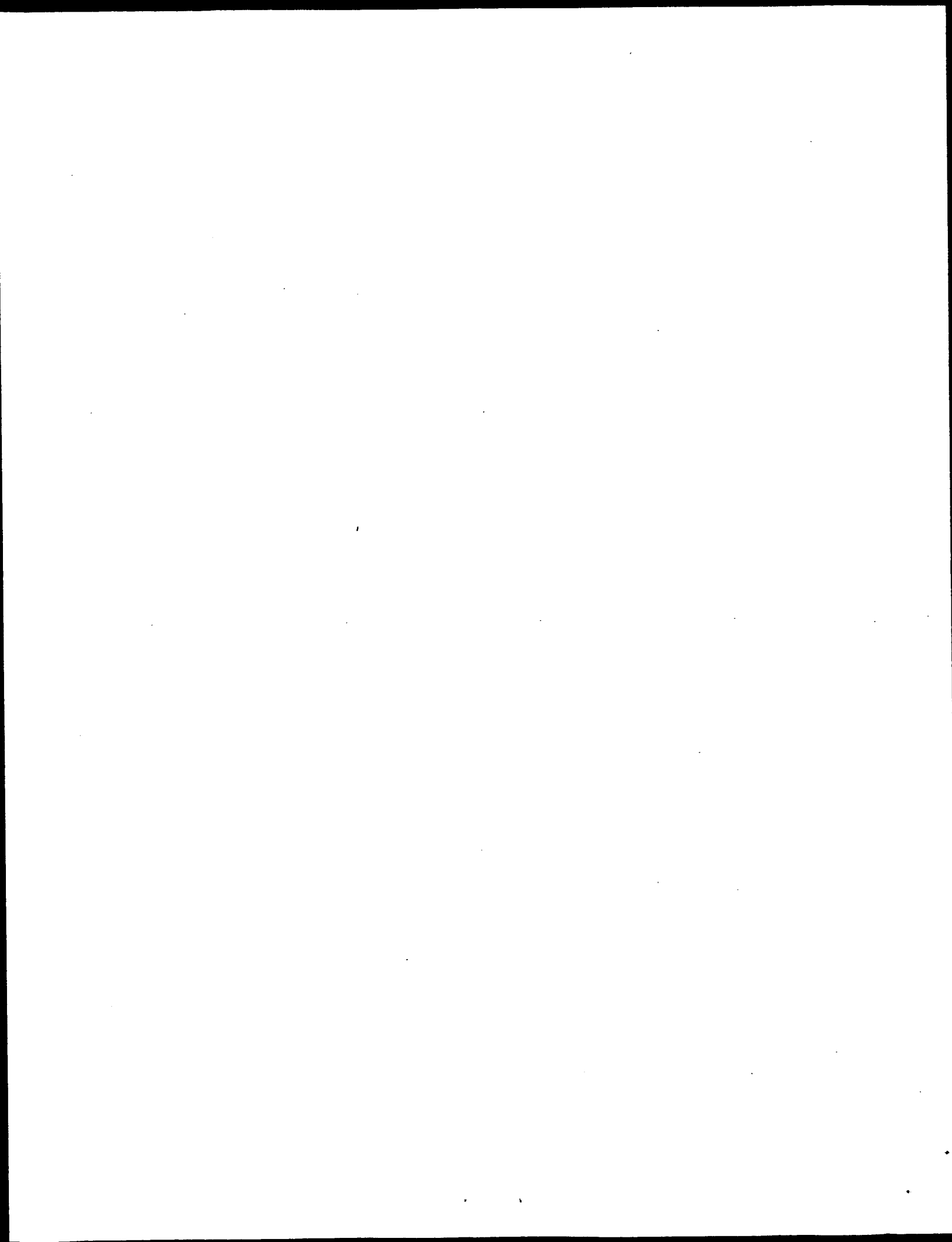
RESULT 14
 ABB60155
 ID ABB60155 standard; Protein; 417 AA.
 AC ABB60155;
 XX
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7257.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 PI WPI; 2001-656860/75.
 XX
 DR N-PSDB; ABL04258.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 7257; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

QY 2 CPXEEKNDRICTNCC 18
 Db 381 CPXGDGSKDLVCFRC 397

RESULT 15
 AAG51616
 ID AAG51616 standard; Protein; 609 AA.
 XX
 AC AAG51616;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 65528.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.

Search completed: December 17, 2002, 10:23:20
Job time : 36 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2002, 10:23:23 ; Search time 15 Seconds
(without alignments)
45.115 Million cell updates/sec

Title: US-09-812-502A-16
Perfect score: 120
Sequence: 1 XCPXXEKKNDRICTNCCAGXKG 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	95.8	368	3	US-08-454-295-3
2	115	95.8	368	4	US-09-431-500A-3
3	115	95.8	368	4	US-09-431-498-3
4	115	95.8	368	4	US-09-431-499-3
5	113	94.2	23	4	US-09-431-500A-16
6	74	61.7	13	3	US-08-454-295-11
7	74	61.7	13	4	US-09-431-500A-11
8	74	61.7	13	4	US-09-431-498-11
9	74	61.7	13	4	US-09-431-499-11
10	74	61.7	58	3	US-08-454-295-5
11	74	61.7	58	3	US-08-454-295-6
12	74	61.7	58	4	US-09-431-500A-5
13	74	61.7	58	4	US-09-431-500A-6
14	74	61.7	58	4	US-09-431-498-5
15	74	61.7	58	4	US-09-431-498-6
16	74	61.7	58	4	US-09-431-499-5
17	74	61.7	58	4	US-09-431-499-6
18	73	60.8	13	4	US-08-454-295-12
19	73	60.8	13	4	US-09-431-500A-12
20	73	60.8	13	4	US-09-431-498-12
21	73	60.8	13	4	US-09-431-499-12
22	73	60.8	54	3	US-08-454-295-10
23	73	60.8	54	4	US-09-431-500A-10
24	73	60.8	54	4	US-09-431-498-10
25	73	60.8	54	4	US-09-431-499-10
26	73	60.8	58	3	US-08-454-295-7
27	73	60.8	58	3	US-08-454-295-8

28	73	60.8	58	3	US-08-454-295-9	Sequence 9, Appli
29	73	60.8	58	4	US-09-431-500A-7	Sequence 7, Appli
30	73	60.8	58	4	US-09-431-500A-8	Sequence 8, Appli
31	73	60.8	58	4	US-09-431-500A-9	Sequence 9, Appli
32	73	60.8	58	4	US-09-431-498-7	Sequence 7, Appli
33	73	60.8	58	4	US-09-431-498-8	Sequence 8, Appli
34	73	60.8	58	4	US-09-431-498-9	Sequence 9, Appli
35	73	60.8	58	4	US-09-431-499-7	Sequence 7, Appli
36	73	60.8	58	4	US-09-431-499-8	Sequence 8, Appli
37	73	60.8	58	4	US-09-431-499-9	Sequence 9, Appli
38	49	40.8	3443	2	US-08-416-603-2	Sequence 2, Appli
39	47	39.2	148	4	US-09-411-722-2	Sequence 1, Appli
40	47	39.2	176	4	US-09-411-722-1	Sequence 2, Appli
41	47	39.2	2508	4	US-09-627-650B-7	Sequence 7, Appli
42	47	39.2	2508	4	US-09-436-063C-7	Sequence 7, Appli
43	47	39.2	2544	4	US-09-627-650B-3	Sequence 3, Appli
44	47	39.2	2544	4	US-09-436-063C-3	Sequence 3, Appli
45	47	39.2	2601	4	US-09-627-650B-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-454-295-3
; Sequence 3, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-295-3

Query Match 95.8%; Score 115; DB 3; Length 368;
Best Local Similarity 86.4%; Pred. No. 3.3e-08;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CFXEKKNDRICTNCCAGXKG 23

Db 16 CPRSEKNDRICTNCCAGTKG 37

RESULT 2

US-09-431-500A-3

Sequence 3, Application US/09431500A

Patent No. 6261821

GENERAL INFORMATION:

APPLICANT: Anderson, Marilyn A.

APPLICANT: Atkinson, Angela H.

APPLICANT: Heath, Robyn L.

APPLICANT: Clarke, Adrienne E.

TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC

TITLE OF INVENTION: SEQUENCES ENCODING SAME

FILE REFERENCE: 9748B

CURRENT APPLICATION NUMBER: US/09/431,500A

CURRENT FILING DATE: 1999-11-01

PRIOR APPLICATION NUMBER: 08/454,295

PRIOR FILING DATE: 1995-09-01

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 3

LENGTH: 368

TYPE: PRT

ORGANISM: Nicotiana glauca

US-09-431-500A-3

Query Match

95.8%; Score 115; DB 4; Length 368;

Best Local Similarity 86.4%;

Pred. No. 3.3e-08;

Matches 19; Conservative 0;

Mismatches 3; Indels 0; Gaps 0;

QY 2 CPXSEKNDRICTNCCAGTKG 23

DB 16 CPXSEKNDRICTNCCAGTKG 37

RESULT 3

US-09-431-498-3

Sequence 3, Application US/09431498

Patent No. 6440727

GENERAL INFORMATION:

APPLICANT: Anderson, Marilyn A.

APPLICANT: Atkinson, Angela H.

APPLICANT: Heath, Robyn L.

APPLICANT: Clarke, Adrienne E.

TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC

TITLE OF INVENTION: SEQUENCES ENCODING SAME

FILE REFERENCE: 9748B

CURRENT APPLICATION NUMBER: US/09/431,498

CURRENT FILING DATE: 1999-11-01

PRIOR APPLICATION NUMBER: 08/454,295

PRIOR FILING DATE: 1995-09-01

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patentin Ver. 2.1

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-431-498-3

Query Match

95.8%; Score 115; DB 4; Length 368;

Best Local Similarity 86.4%;

Pred. No. 3.3e-08;

Matches 19; Conservative 0;

Mismatches 3; Indels 0; Gaps 0;

QY 2 CPXSEKNDRICTNCCAGTKG 23

DB 16 CPXSEKNDRICTNCCAGTKG 37

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-431-498-3

Query Match

95.8%; Score 115; DB 4; Length 368;

Best Local Similarity 86.4%;

Pred. No. 3.3e-08;

Matches 19; Conservative 0;

Mismatches 3; Indels 0; Gaps 0;

QY 2 CPXSEKNDRICTNCCAGTKG 23

DB 16 CPXSEKNDRICTNCCAGTKG 37

RESULT 4

US-09-431-499-3

Sequence 3, Application US/09431499

Patent No. 6451573

GENERAL INFORMATION:

APPLICANT: Anderson, Marilyn A.

APPLICANT: Atkinson, Angela H.

APPLICANT: Heath, Robyn L.

APPLICANT: Clarke, Adrienne E.

TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF

TITLE OF INVENTION: AND GENETIC

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: United States of America

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/431,499

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/454,295

FILING DATE: 01-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9748

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-431-499-3

Query Match

95.8%; Score 115; DB 4; Length 368;

Best Local Similarity 86.4%;

Pred. No. 3.3e-08;

Matches 19; Conservative 0;

Mismatches 3; Indels 0; Gaps 0;

QY 2 CPXSEKNDRICTNCCAGTKG 23

DB 16 CPXSEKNDRICTNCCAGTKG 37


```

RESULT 5
US-09-431-500A-16
; Sequence 16, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; NAME/KEY: UNSURE
; LOCATION: (1)
; OTHER INFORMATION: Xaa is Ile or Val
; NAME/KEY: UNSURE
; LOCATION: (4)
; OTHER INFORMATION: Xaa is Arg or Leu
; NAME/KEY: UNSURE
; LOCATION: (5)
; OTHER INFORMATION: Xaa is Ser or Ala
; NAME/KEY: UNSURE
; LOCATION: (21)
; * OTHER INFORMATION: Xaa is Thr or Lys
US-09-431-500A-16

Query Match          94.2%; Score 113; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 CPXEEKKNDRICTNCCAGXKG 23
    |||||
Db   2 CPXEEKKNDRICTNCCAGXKG 23
    |||||

RESULT 6
US-08-454-295-11
; Sequence 11, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295

```

```

; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-454-295-11

Query Match          61.7%; Score 74; DB 3; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00048;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  11 DRICTNCCAGXKG 23
    |||||
Db   1 DRICTNCCAGTKG 13
    |||||

RESULT 7
US-09-431-500A-11
; Sequence 11, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/431,500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-431-500A-11

Query Match          61.7%; Score 74; DB 4; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00048;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  11 DRICTNCCAGXKG 23
    |||||
Db   1 DRICTNCCAGTKG 13
    |||||

RESULT 8
US-09-431-498-11
; Sequence 11, Application US/09431498
; Patent No. 640727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,498
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/454,295
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-431-498-11
Query Match 61.7%; Score 74; DB 4; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00048;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 11 DRICNCCAGKKG 23
Db 1 DRICNCCAGKKG 13
RESULT 9
US-09-431-499-11
Sequence 11, Application US/09431499
Patent No. 6451573
GENERAL INFORMATION:
APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,499
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/454,295
FILING DATE: 01-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-431-499-11
Query Match 61.7%; Score 74; DB 4; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00048;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 11 DRICNCCAGKKG 23
Db 1 DRICNCCAGKKG 13
RESULT 10
US-08-454-295-5
Sequence 5, Application US/08454295
Patent No. 6031087
GENERAL INFORMATION:
APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,295
FILING DATE: 01-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-454-295-5
Query Match 61.7%; Score 74; DB 3; Length 58;

```
Best Local Similarity 92.3%; Pred. NO. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 DR1CTNCCAGXKG 23
Db 1 DR1CTNCCAGTKG 13

RESULT 11
US-08-454-295-6
; Sequence 6, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-295-6

Query Match 61.7%; Score 74; DB 3; Length 58;
Best Local Similarity 92.3%; Pred. NO. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 DR1CTNCCAGXKG 23
Db 1 DR1CTNCCAGTKG 13

RESULT 12
US-09-431-500A-5
; Sequence 5, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
```

```
; CURRENT APPLICATION NUMBER: US/09/431.500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-431-500A-5

Query Match 61.7%; Score 74; DB 4; Length 58;
Best Local Similarity 92.3%; Pred. NO. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 DR1CTNCCAGXKG 23
Db 1 DR1CTNCCAGTKG 13

RESULT 13
US-09-431-500A-6
; Sequence 6, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/431.500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-431-500A-6

Query Match 61.7%; Score 74; DB 4; Length 58;
Best Local Similarity 92.3%; Pred. NO. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 DR1CTNCCAGXKG 23
Db 1 DR1CTNCCAGTKG 13

RESULT 14
US-09-431-498-5
; Sequence 5, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
```

Tue Dec 17 14:06:01 2002

us-09-812-502a-16.ra1

Page 6-

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,498
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/454,295
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-431-498-5

Query Match 61.7%; Score 74; DB 4; Length 58;
Best Local Similarity 92.3%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0;

QY 11 DRICNCCAGXKG 23
DB 1 DRICNCCAGXKG 13

RESULT 15
US-09-431-498-6
Sequence 6, Application US/09431498
Patent No. 6440727
GENERAL INFORMATION:
APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,498
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/454,295
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-431-498-6

Query Match 61.7%; Score 74; DB 4; Length 58;
Best Local Similarity 92.3%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0;

QY 11 DRICNCCAGXKG 23
DB 1 DRICNCCAGXKG 13

Search completed: December 17, 2002, 10:27:21
Job time : 15 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2002, 10:23:44 ; Search time 11 Seconds
(without alignments)
34.839 Million cell updates/sec

Title: US-09-812-502A-16

Perfect score: 120

Sequence: 1 XCPXEEKNDRICTNCCAGK 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 16662342 residues

Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
- 9: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	39.2	148	10	US-09-855-266A-2
2	47	39.2	176	10	US-09-855-266A-1
3	46	38.3	3907	9	US-10-029-217A-24
4	45	37.5	1200	10	US-09-826-508-3
5	45	37.5	1300	12	US-10-052-586-269
6	45	37.5	2211	9	US-10-096-961-1
7	44.5	37.1	1609	10	US-09-938-275-11
8	44	36.7	873	9	US-10-167-264-2
9	44	36.7	1009	10	US-09-898-570-16
10	44	36.7	4440	12	US-10-052-586-525
11	44	36.7	4679	10	US-09-804-898-2
12	43	35.8	1743	12	US-10-052-586-451
13	42	35.0	35	8	US-08-969-137-2
14	42	35.0	153	10	US-09-801-368-434
15	42	35.0	162	10	US-09-871-388-10
16	41	34.2	110	10	US-09-864-761-46494
17	41	34.2	110	10	US-09-864-761-47545
18	41	34.2	114	10	US-09-925-300-1175
19	41	34.2	564	10	US-09-764-864-1245

20	41	34.2	749	10	US-09-871-388-8	Sequence 8, Appli
21	41	34.2	975	10	US-09-886-055-431	Sequence 431, App
22	41	34.2	1400	10	US-09-879-957-37	Sequence 37, Appl
23	41	34.2	1435	12	US-10-052-586-581	Sequence 581, App
24	41	34.2	1497	10	US-09-060-854B-2	Sequence 2, Appli
25	41	34.2	1894	12	US-10-052-586-97	Sequence 2, Appli
26	40	33.3	410	10	US-09-983-931-2	Sequence 2, Appli
27	40	33.3	969	10	US-09-969-515-10	Sequence 10, Appli
28	40	33.3	980	10	US-09-969-515-4	Sequence 4, Appli
29	40	33.3	1002	9	US-09-988-117-3	Sequence 3, Appli
30	40	33.3	1002	10	US-09-812-471-3	Sequence 3, Appli
31	40	33.3	1002	10	US-09-812-633-3	Sequence 3, Appli
32	40	33.3	1005	9	US-09-988-117-1	Sequence 1, Appli
33	40	33.3	1005	10	US-09-812-471-1	Sequence 1, Appli
34	40	33.3	1005	10	US-09-812-633-1	Sequence 1, Appli
35	40	33.3	1213	10	US-09-969-515-8	Sequence 8, Appli
36	40	33.3	1224	10	US-09-969-515-2	Sequence 2, Appli
37	40	33.3	1629	10	US-09-972-467-2	Sequence 2, Appli
38	40	33.3	1882	10	US-09-918-171A-13	Sequence 13, Appli
39	40	33.3	1907	10	US-09-938-330-25	Sequence 25, Appli
40	40	33.3	3084	10	US-09-938-275-4	Sequence 4, Appli
41	39.5	32.9	383	10	US-09-205-658-105	Sequence 105, App
42	39.5	32.9	383	10	US-09-844-353A-105	Sequence 105, App
43	39.5	32.9	1607	10	US-09-938-275-10	Sequence 10, Appli
44	39	32.5	30	10	US-09-864-761-41978	Sequence 41978, A
45	39	32.5	90	10	US-09-864-761-35453	Sequence 35453, A

ALIGNMENTS

RESULT 1
US-09-855-266A-2
; Sequence 2, Application US/09855266A
; Patent No. US20020128435A1
; GENERAL INFORMATION:
; APPLICANT: Kimura, Naoki
; APPLICANT: Toyoshima, Tomoko
; TITLE OF INVENTION: NOVEL SECRETORY MEMBRANE PROTEIN
; FILE REFERENCE: 06501-040002
; CURRENT APPLICATION NUMBER: US/09/855,266A
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 09/411,722
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/JP98/01511
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: JP 9/099653
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-855-266A-2

Query Match 39.2%; Score 47; DB 10; Length 148;
Best Local Similarity 42.1%; Pred. No. 8.1;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXEEKNDRICTNCCAG 20
DB 10 CPDGEYQSDVCKCTCPG 28

RESULT 2
US-09-855-266A-1
; Sequence 1, Application US/09855266A
; Patent No. US20020128435A1
; GENERAL INFORMATION:
; APPLICANT: Kimura, Naoki
; APPLICANT: Toyoshima, Tomoko
; TITLE OF INVENTION: NOVEL SECRETORY MEMBRANE PROTEIN

FILE REFERENCE: 06501-040002
CURRENT APPLICATION NUMBER: US/09/855,266A
CURRENT FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: US 09/411,722
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: PCT/JP98/01511
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: JP 9/099653
PRIOR FILING DATE: 1997-04-01
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 176
TYPE: PRT
ORGANISM: Mus musculus
US-09-855-266A-1

Query Match 39.2%; Score 47; DB 10; Length 176;
Best Local Similarity 42.1%; Pred. No. 9.4;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 2 CPXXEKKNDRICTNCCAG 20
Db 38 CPDGEYQSNVCKCTCPSG 56

RESULT 3
US-10-029-217A-24
Sequence 24, Application US/10029217A
Patent No. US20020164735A1
GENERAL INFORMATION:
APPLICANT: OLSON, ERIC N.
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
FILE REFERENCE: UTS:69505
CURRENT APPLICATION NUMBER: US/10/029,217A
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: 60/257,761
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 3907
TYPE: PRT
ORGANISM: Mus musculus
US-10-029-217A-24

Query Match 38.3%; Score 46; DB 9; Length 3907;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 CTNCCAGXKG 23
Db 3652 CTGCCAGAG 3661

RESULT 4
US-09-826-508-3
Sequence 3, Application US/09826508
Patent No. US2001002509A1
GENERAL INFORMATION:
APPLICANT: Nabil Elshourbagy
TITLE OF INVENTION: G-Protein-Coupled Receptor Polypeptides
FILE REFERENCE: GP-70744USB
CURRENT APPLICATION NUMBER: US/09/826,508
CURRENT FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1200

TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-826-508-3

Query Match 37.5%; Score 45; DB 10; Length 1200;
Best Local Similarity 70.0%; Pred. No. 87;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 CTNCCAGXKG 23
Db 602 CTGCCAGCTG 611

RESULT 5
US-10-052-586-269
Sequence 269, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065111
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870

1 PRIOR FILING DATE: 1997-12-17
2 PRIOR APPLICATION NUMBER: 60/068017
3 PRIOR FILING DATE: 1997-12-18
4 PRIOR APPLICATION NUMBER: 60/077450
5 PRIOR FILING DATE: 1998-03-10
6 PRIOR APPLICATION NUMBER: 60/077632
7 PRIOR FILING DATE: 1998-03-11
8 PRIOR APPLICATION NUMBER: 60/077649
9 PRIOR FILING DATE: 1998-03-11
10 PRIOR APPLICATION NUMBER: 60/078886
11 PRIOR FILING DATE: 1998-03-20
12 PRIOR APPLICATION NUMBER: 60/078939
13 PRIOR FILING DATE: 1998-03-20
14 PRIOR APPLICATION NUMBER: 60/079664
15 PRIOR FILING DATE: 1998-03-27
16 PRIOR APPLICATION NUMBER: 60/079786
17 PRIOR FILING DATE: 1998-03-27
18 PRIOR APPLICATION NUMBER: 60/080107
19 PRIOR FILING DATE: 1998-03-31
20 PRIOR APPLICATION NUMBER: 60/080194
21 PRIOR FILING DATE: 1998-03-31
22 PRIOR APPLICATION NUMBER: 60/080327
23 PRIOR FILING DATE: 1998-04-01
24 PRIOR APPLICATION NUMBER: 60/080333
25 PRIOR FILING DATE: 1998-04-01
26 PRIOR APPLICATION NUMBER: 60/081049
27 PRIOR FILING DATE: 1998-04-08
28 PRIOR APPLICATION NUMBER: 60/081070
29 PRIOR FILING DATE: 1998-04-08
30 PRIOR APPLICATION NUMBER: 60/081195
31 PRIOR FILING DATE: 1998-04-09
32 PRIOR APPLICATION NUMBER: 60/081838
33 PRIOR FILING DATE: 1998-04-15
34 PRIOR APPLICATION NUMBER: 60/082568
35 PRIOR FILING DATE: 1998-04-21
36 PRIOR APPLICATION NUMBER: 60/082569
37 PRIOR FILING DATE: 1998-04-21
38 PRIOR APPLICATION NUMBER: 60/082704
39 PRIOR FILING DATE: 1998-04-22
40 PRIOR APPLICATION NUMBER: 60/082797
41 PRIOR FILING DATE: 1998-04-22
42 PRIOR APPLICATION NUMBER: 60/083322
43 PRIOR FILING DATE: 1998-04-28
44 PRIOR APPLICATION NUMBER: 60/083495
45 PRIOR FILING DATE: 1998-04-29
46 PRIOR APPLICATION NUMBER: 60/083496
47 PRIOR FILING DATE: 1998-04-29
48 PRIOR APPLICATION NUMBER: 60/083499
49 PRIOR FILING DATE: 1998-04-29
50 PRIOR APPLICATION NUMBER: 60/083559
51 PRIOR FILING DATE: 1998-04-29
52 PRIOR APPLICATION NUMBER: 60/084366
53 PRIOR FILING DATE: 1998-05-05
54 PRIOR APPLICATION NUMBER: 60/084414
55 PRIOR FILING DATE: 1998-05-06
56 PRIOR APPLICATION NUMBER: 60/084639
57 PRIOR FILING DATE: 1998-05-07
58 PRIOR APPLICATION NUMBER: 60/084640
59 PRIOR FILING DATE: 1998-05-07
60 PRIOR APPLICATION NUMBER: 60/084643
61 PRIOR FILING DATE: 1998-05-07
62 PRIOR APPLICATION NUMBER: 60/085573
63 PRIOR FILING DATE: 1998-05-15
64 PRIOR APPLICATION NUMBER: 60/085579
65 PRIOR FILING DATE: 1998-05-15
66 PRIOR APPLICATION NUMBER: 60/085580
67 PRIOR FILING DATE: 1998-05-15
68 PRIOR APPLICATION NUMBER: 60/085582
69 PRIOR FILING DATE: 1998-05-15
70 PRIOR APPLICATION NUMBER: 60/085700
71 PRIOR FILING DATE: 1998-05-15
72 PRIOR APPLICATION NUMBER: 60/086023
73 PRIOR FILING DATE: 1998-05-18

1 PRIOR APPLICATION NUMBER: 60/086392
2 PRIOR FILING DATE: 1998-05-22
3 PRIOR APPLICATION NUMBER: 60/086486
4 PRIOR FILING DATE: 1998-05-22
5 PRIOR APPLICATION NUMBER: 60/087098
6 PRIOR FILING DATE: 1998-05-28
7 PRIOR APPLICATION NUMBER: 60/087208
8 PRIOR FILING DATE: 1998-05-28
9 PRIOR APPLICATION NUMBER: 60/087609
10 PRIOR FILING DATE: 1998-06-02
11 PRIOR APPLICATION NUMBER: 60/087759
12 PRIOR FILING DATE: 1998-06-02
13 PRIOR APPLICATION NUMBER: 60/087827
14 PRIOR FILING DATE: 1998-06-03
15 PRIOR APPLICATION NUMBER: 60/088025
16 PRIOR FILING DATE: 1998-06-04
17 PRIOR APPLICATION NUMBER: 60/088028
18 PRIOR FILING DATE: 1998-06-04
19 PRIOR APPLICATION NUMBER: 60/088029
20 PRIOR FILING DATE: 1998-06-04
21 PRIOR APPLICATION NUMBER: 60/088033
22 PRIOR FILING DATE: 1998-06-04
23 PRIOR APPLICATION NUMBER: 60/088167
24 PRIOR FILING DATE: 1998-06-05
25 PRIOR APPLICATION NUMBER: 60/088202
26 PRIOR FILING DATE: 1998-06-05
27 PRIOR APPLICATION NUMBER: 60/088212
28 PRIOR FILING DATE: 1998-06-05
29 PRIOR APPLICATION NUMBER: 60/088217
30 PRIOR FILING DATE: 1998-06-05
31 PRIOR APPLICATION NUMBER: 60/088326
32 PRIOR FILING DATE: 1998-06-04
33 PRIOR APPLICATION NUMBER: 60/088655
34 PRIOR FILING DATE: 1998-06-09
35 PRIOR APPLICATION NUMBER: 60/088722
36 PRIOR FILING DATE: 1998-06-10
37 PRIOR APPLICATION NUMBER: 60/088738
38 PRIOR FILING DATE: 1998-06-10
39 PRIOR APPLICATION NUMBER: 60/088740
40 PRIOR FILING DATE: 1998-06-10
41 PRIOR APPLICATION NUMBER: 60/088811
42 PRIOR FILING DATE: 1998-06-10
43 PRIOR APPLICATION NUMBER: 60/088824
44 PRIOR FILING DATE: 1998-06-10
45 PRIOR APPLICATION NUMBER: 60/088825
46 PRIOR FILING DATE: 1998-06-10
47 PRIOR APPLICATION NUMBER: 60/088826
48 PRIOR FILING DATE: 1998-06-10
49 PRIOR APPLICATION NUMBER: 60/088861
50 PRIOR FILING DATE: 1998-06-11
51 PRIOR APPLICATION NUMBER: 60/088863
52 PRIOR FILING DATE: 1998-06-11
53 PRIOR APPLICATION NUMBER: 60/088876
54 PRIOR FILING DATE: 1998-06-11
55 PRIOR APPLICATION NUMBER: 60/089090
56 PRIOR FILING DATE: 1998-06-12
57 PRIOR APPLICATION NUMBER: 60/089105
58 PRIOR FILING DATE: 1998-06-12
59 PRIOR APPLICATION NUMBER: 60/089512
60 PRIOR FILING DATE: 1998-06-16
61 PRIOR APPLICATION NUMBER: 60/089514
62 PRIOR FILING DATE: 1998-06-16
63 PRIOR APPLICATION NUMBER: 60/089538
64 PRIOR FILING DATE: 1998-06-17
65 PRIOR APPLICATION NUMBER: 60/089598
66 PRIOR FILING DATE: 1998-06-17
67 PRIOR APPLICATION NUMBER: 60/089653
68 PRIOR FILING DATE: 1998-06-17
69 PRIOR APPLICATION NUMBER: 60/089908

Query Match 37.5%; Score 45; DB 12; Length 1300;
Best Local Similarity 70.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 14 CTNCCAGXKG 23
 DB 918 CTGCCAGTCG 927

RESULT 6

US-10-096-961-1
 ; Sequence 1, Application US/10096961
 ; Patent No. US2002015572A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUEGLER, Karl et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
 ; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: CLO00849D1V
 ; CURRENT APPLICATION NUMBER: US/10/096,961
 ; CURRENT FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: 60/232,632
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 09/738,884
 ; PRIOR FILING DATE: 2000-12-18
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2211
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-10-096-961-1

Query Match 37.5%; Score 45; DB 9; Length 2211;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 14 CTNCCAGXKG 23
 DB 1680 CTGCCAGACG 1689

RESULT 7
 US-09-938-275-11
 ; Sequence 11, Application US/09938275
 ; Patent No. US2002011309A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gerardo Castillo
 ; APPLICANT: Alan Snow
 ; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
 ; TITLE OF INVENTION: of Laminin and Laminin-Derived Protein Fragments
 ; FILE REFERENCE: PROTEO.P03
 ; CURRENT APPLICATION NUMBER: US/09/938,275
 ; CURRENT FILING DATE: 2001-08-16
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 1609
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: Swissprot P11047
 ; DATABASE ENTRY DATE: 1991-11-01
 ; US-09-938-275-11

Query Match 37.1%; Score 44.5; DB 10; Length 1609;
 Best Local Similarity 33.3%; Pred. No. 1.3e+02;
 Matches 9; Conservative 2; Mismatches 11; Indels 5; Gaps 1;

OY 2 CPXHE---KANDRICNCCAGXKG 23
 DB 775 CPGSSCAVVPKKEVCTNCPGTGTG 801

RESULT 8
 US-10-167-264-2

; Sequence 2, Application US/10167264
 ; Publication No. US20020182182A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wilson, James M.
 ; APPLICANT: Kozarsky, Karen F.
 ; APPLICANT: Straus, Jerome F.

TITLE OF INVENTION: Methods and Compositions for Gene
 Therapy for the Treatment of Defects in Lipoprotein
 Metabolism

NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Howson and Howson
 STREET: Spring House Corporate Cntr., PO Box 457
 CITY: Spring House
 STATE: Pennsylvania
 COUNTRY: USA
 ZIP: 19477

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/167,264
 FILING DATE: 10-Jun-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/894,489
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 08/393,734
 FILING DATE: 24-FEB-1995

ATTORNEY/AGENT INFORMATION:
 NAME: Bak, Mary E.
 REGISTRATION NUMBER: 31,215
 REFERENCE/DOCKET NUMBER: GAVPN.009CIP1USA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-540-9200
 TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 873 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 ; US-10-167-264-2

Query Match 36.7%; Score 44; DB 9; Length 873;
 Best Local Similarity 36.4%; Pred. No. 90;
 Matches 8; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 2 CPXHEKNDRICTNCCAGXKG 23
 DB 394 CGDIDCQNPICISQICINLKG 415

RESULT 9
 US-09-898-570-16
 ; Sequence 16, Application US/09898570
 ; Patent No. US20020123612A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GERLACH, VALERIE L.
 ; APPLICANT: ELLERMAN, KAREN
 ; APPLICANT: MACDOUGALL, JOHN R.
 ; APPLICANT: SMITHSON, GLENDA
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
 ; TITLE OF INVENTION: METHODS OF USING THE SAME
 ; FILE REFERENCE: 15966-776C1P
 ; CURRENT APPLICATION NUMBER: US/09/898,570
 ; CURRENT FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/198,293
 ; PRIOR FILING DATE: 2000-04-19
 ; PRIOR APPLICATION NUMBER: 60/198,645


```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: CG55096-04
US-09-898-570-16

Query Match 35.7%; Score 44; DB 10; Length 1009;
Best Local Similarity 31.8%; Pred. No. 1e-02;
Matches 7; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Qy' 2 CPXBEKNDRICTNCCAGXKG 23
Db 332 CODIDERSFDRDCHICVNTG 353

RESULT 10
US-10-052-586-525
; Sequence 525, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
```

PRIOR APPLICATION NUMBER: 60/083436
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083439
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086023
 PRIOR FILING DATE: 1998-05-18
 PRIOR APPLICATION NUMBER: 60/086392
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086486
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087098
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087208
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088722
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088740
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088811
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088825

PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088863
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089090
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089908

Query Match 36.7%; Score 44; DB 12; Length 4440;
 Best Local Similarity 70.0%; Pred. No. 3.6e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CTNCCAGXKG 23
 DB 550 CTNCCAGAG 559

RESULT 11
 US-09-804-898-2
 Sequence 2, Application US/09804898
 Patent No. US20020045264A1
 GENERAL INFORMATION:
 APPLICANT: DURING, MATTHEW
 APPLICANT: XIAO, WEIDONG
 TITLE OF INVENTION: PRODUCTION OF CHIMERIC CAPSID VECTORS
 FILE REFERENCE: 102182-14
 CURRENT APPLICATION NUMBER: US/09/804,898
 CURRENT FILING DATE: 2001-03-13
 PRIOR FILING DATE: 2000-03-14
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 2
 LENGTH: 4679
 TYPE: PRT
 ORGANISM: adeno-associated virus 2
 US-09-804-898-2

Query Match 36.7%; Score 44; DB 10; Length 4679;
 Best Local Similarity 70.0%; Pred. No. 3.7e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 CTNCCAGXKG 23
 DB 2556 CTNCCAGGCG 2565

RESULT 12
 US-10-052-586-451
 Sequence 451, Application US/10052586
 Patent No. US20020127584A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-17

Query Match 35.8%; Score 43; DB 12; Length 1743;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 CTNCCAGK 23
Db 1235 CTTCCTGATG 1244

RESULT 13
US-08-969-137-2
Sequence 2, Application US/08969137
Patent No. US20010018207A1
GENERAL INFORMATION:
APPLICANT: KANDEL, ERIC
APPLICANT: MAYFORD, MARK
TITLE OF INVENTION: DNA REGULATORY ELEMENT FOR THE
TITLE OF INVENTION: EXPRESSION OF TRANSGENES IN NEURONS OF THE MOUSE FOREBRAIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER AND DUNHAM
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/969,137
FILING DATE: 12-NOV-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52776
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)2276821
TELEFAX: (212)3910525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-969-137-2

Query Match 35.0%; Score 42; DB 8; Length 35;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 CTNCCAG 20
Db 7 CTGCCAG 13

RESULT 14
US-09-801-368-434
Sequence 434, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: Patent in version 3.0
SEQ ID NO 434
LENGTH: 153
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-434

Query Match 35.0%; Score 42; DB 10; Length 153;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 KNDRICTNC 17
 Db 132 QNDRLCYNC 140

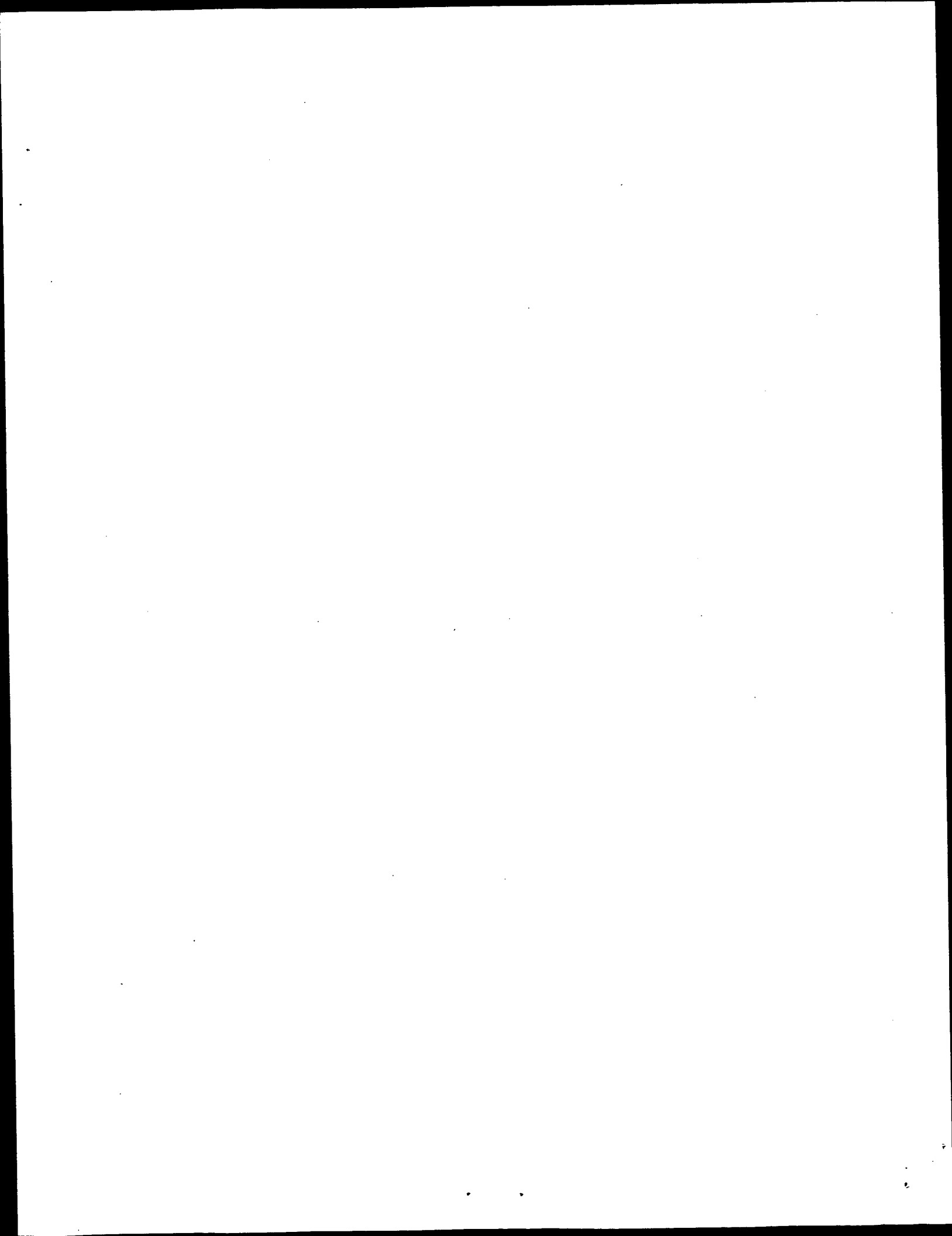
RESULT 15

US-09-871-388-10
 ; Sequence 10, Application US/09871388
 ; Patent No. US20020127621A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rubin, Gerald M.
 ; Pan, Duojia
 ; Rooke, Jenny
 ; Yavari, Reza
 ; Xu, Tian
 ;
 ; TITLE OF INVENTION: KUZ: A No. US20020127621A1e1 Family of Metalloproteases
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 BUSH STREET, SUITE 3200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94104
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/871,388
 ; FILING DATE: 31-May-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/937,931
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OSMAN, RICHARD A
 ; REGISTRATION NUMBER: 16,627
 ; REFERENCE/DOCKET NUMBER: B97-081
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 343-4341
 ; TELEFAX: (415) 343-4342
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 162 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 ; US-09-871-388-10

Query Match 35.0%; Score 42; DB 10; Length 162;
 Best Local Similarity 35.3%; Pred. NO. 39;
 Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCC 18
 Db 108 CGSTDEKDDKELCHVCC 124

Search completed: December 17, 2002, 10:27:39
 Job time : 12 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2002, 10:22:28 ; Search time 16 Seconds
(without alignments)
138.193 Million cell updates/sec

Title: US-09-812-502A-16

Perfect score: 120

Sequence: 1 XCPXXEKKNDRICTNCCAGXKG 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	95.8	397	JQ2153	proteinase inhibit
2	85	70.8	204	T08072	proteinase inhibit
3	82	68.3	223	S43338	proteinase inhibit
4	78	65.0	153	XKPOC1	proteinase inhibit
5	78	65.0	158	T07597	proteinase inhibit
6	75	62.5	201	T07011	proteinase inhibit
7	74.5	62.1	146	S72492	proteinase inhibit
8	74	61.7	53	JQ2269	probable proteinase
9	74	61.7	197	S56662	trypsin inhibitor
10	69	57.5	147	S24973	proteinase inhibit
11	69	57.5	154	S43105	proteinase inhibit
12	67	55.8	51	XKPO2	proteinase inhibit
13	66	55.0	148	B24048	proteinase inhibit
14	58	48.3	45	XKPO2A	proteinase inhibit
15	51	42.5	52	T1501	proteinase inhibit
16	50	41.7	40	XKPO2B	proteinase inhibit
17	49	40.8	919	S45889	probable regulator
18	47	39.2	762	E96593	probable phospholi
19	47	39.2	869	JC4858	VLDL receptor prec
20	46	38.3	604	C87818	protein gih-1 limp
21	46	38.3	604	T15132	ATP-dependent RNA
22	46	38.3	707	A48686	probable RNA helic
23	46	38.3	1906	S68235	myosin-light-chain
24	45	37.5	164	A96500	hypothetical prote
25	45	37.5	383	H84156	hypothetical prote
26	45	37.5	761	B64506	amino transferase
27	44.5	37.1	1609	1MHUB2	DNA topoisomerase
28	44	36.7	77	I51647	laminin gamma-1 ch
29	44	36.7	92	A72242	tenascin - African ferredoxin - Therm

30 44 36.7 321 2 AE0304 probable dioxygena
31 44 36.7 385 2 S42378 hypothetical prote
32 44 36.7 521 2 H69317 conserved hypothet
33 44 36.7 689 2 T08988 cadmium-transporti
34 44 36.7 711 2 A85352 cadmium-transporti
35 44 36.7 863 1 S51789 VLDL receptor prec
36 44 36.7 873 1 A49729 VLDL receptor prec
37 44 36.7 873 1 QRR8VD VLDL receptor prec
38 44 36.7 873 1 I48952 VLDL receptor prec
39 44 36.7 1356 2 A45445 janusin precursor
40 44 36.7 1810 1 A32230 tenascin precursor
41 44 36.7 2510 2 T28160 hypothetical prote
42 43.5 36.2 63 2 A34905 hypothetical prote
43 43.5 36.2 74 2 T24715 hypothetical prote
44 43.5 36.2 441 2 F71425 hypothetical prote
45 43 35.8 389 2 E86793 hypothetical prote

ALIGNMENTS

RESULT 1

JQ2153

proteinase inhibitor II precursor - Persian tobacco

C:Species: Nicotiana glauca (Persian tobacco)

C>Date: 28-Aug-1985 #sequence revision 07-Oct-1994 #text_change 28-May-1999

C:Accession: JQ2153; PQ0647; S65396

R:Atkinson, A.H.; Heath, R.L.; Simpson, R.J.; Clarke, A.E.; Anderson, M.A.

Plant Cell 5, 203-213, 1993

A>Title: Proteinase inhibitors in Nicotiana glauca stigmas are derived from a precursor p

A:Reference number: JQ2153; MUID:93200805; PMID:6453302

A:Accession: JQ2153

A:Molecule type: mRNA

A:Residues: 1-397 <ATK>

A:Cross-references: GB:U08219; MID:9473590; PIDN:AAA17739.1; PID:9473591

A:Accession: PQ0647

A:Molecule type: protein

A:Residues: 54-66;112-124;170-182;228-240;286-298;344-357 <AT2>

A:Experimental source: stigma, style

R:Heath, R.L.; Barton, P.A.; Simpson, R.J.; Reid, G.E.; Lim, G.; Anderson, M.A.

Eur. J. Biochem. 230, 250-257, 1995

A>Title: Characterization of the protease processing sites in a multidomain proteinase in

A:Reference number: S65396; MUID:95324532; PMID:7601108

A:Accession: S65396

A:Molecule type: protein

A:Residues: 54-227;286-343 <HEA>

C:Comment: This protein contains six similar domains, each with a potential active site.

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: serine proteinase inhibitor

F;1-29/Domain: signal sequence #status predicted <SIG>

F;30-397/Product: proteinase inhibitor II #status predicted <MAT>

F;35-36,93-94/Region: chymotrypsin-specific sites

F;153-154,211-212,269-270,327-328/Region: trypsin-specific sites

Query Match

Best Local Similarity 95.8%; Score 115; DB 2; Length 397;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CPXEXEKKNDRICTNCCAGXKG 23

DB 45 CPRSEKKNDRICTNCCAGTKG 66

RESULT 2

T08072

proteinase inhibitor II precursor - pepper

C:Species: Capsicum annuum (pepper)

C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 08-Oct-1999

C:Accession: T08072

R:Kim, S.H.; Choi, D.S.; Lee, K.W.

submitted to the EMBL Data Library, December 1997

A:Description: Isolation and characterization of wound-induced proteinase inhibitor II c

A:Reference number: Z16334

A:Accession: T08072
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-204 <KIM>
 A:Cross-references: EMBL:AF039398; NID:g2745897; PIDD:AA894771.1; PIDD:g2745898
 C:Genetics:
 A:Gene: Pin2
 A:Note: wound-induced
 C:Superfamily: potato proteinase inhibitor PTI
 C:Keywords: serine proteinase inhibitor

Query Match 70.8%; Score 85; DB 2; Length 204;
 Best Local Similarity 63.6%; Pred. No. 0.00012; Indels 0; Gaps 0;
 Matches 14; Conservative 1; Mismatches 7;

QY 2 CPXXEKKNDRICTNCCAGXKG 23
 DB 84 CPRSEGNMRICITNCCAGXKG 105

RESULT 3

S4338
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S4338
 R:Taylor, B.H.; Young, R.J.; Scheuring, C.F.
 Plant Mol. Biol. 23, 1005-1014, 1993
 A>Title: Induction of a proteinase inhibitor II-class gene by auxin in tomato roots.
 A:Reference number: S4338; MUID:94083553; PMID:7903168
 A:Accession: S4338
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-223 <YVY>
 A:Cross-references: EMBL:L21194; NID:9405581; PIDD:AA16881.1; PIDD:9405582
 C:Superfamily: potato proteinase inhibitor PTI

Query Match 68.3%; Score 82; DB 2; Length 223;
 Best Local Similarity 59.1%; Pred. No. 0.00033; Indels 0; Gaps 0;
 Matches 13; Conservative 3; Mismatches 6;

QY 2 CPXXEKKNDRICTNCCAGXKG 23
 DB 104 CPLETKRVEGLCTNCCAGXKG 125

RESULT 4

XPPOCI
 N:Alternate names: potato chymotrypsin inhibitor I; protease inhibitor II
 C:Species: Solanum tuberosum (potato)
 C:Date: 06-Jul-1982 #sequence_revision 12-Apr-1996 #text_change 20-Apr-2000
 C:Accession: A26584; A01319; A33591
 R:Thornbury, R.W.; An, G.; Cleveland, T.E.; Johnson, R.; Ryan, C.A.
 Proc. Natl. Acad. Sci. U.S.A. 84, 744-748, 1987
 A>Title: Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase
 A:Reference number: A26584
 A:Accession: A26584
 A:Molecule type: DNA
 A:Residues: 1-153 <THO>
 A:Cross-references: GB:M15186; NID:g169488; PIDD:AA33815.1; PIDD:g169489
 R:Haas, G.M.; Herndon, M.A.; Ryan, C.A.; Gentry, L.
 Biochemistry 21, 752-756, 1982
 A>Title: Primary structures of two low molecular weight proteinase inhibitors from potato
 A:Reference number: A90465; MUID:82182863; PMID:7074039
 A:Accession: A01319
 A:Molecule type: protein
 A:Residues: 55-106 <HAS>
 A:Note: Leu-92 is probably the site of interaction with chymotrypsin
 R:Kell, M.; Sanchez-Serrano, J.; Scheil, J.; Willmitzer, L.
 Nucleic Acids Res. 14, 5641-5650, 1986
 A>Title: Primary structure of a proteinase inhibitor II gene from potato (Solanum tuberosum)
 A:Reference number: A23591; MUID:86286579; PMID:3016659

A:Accession: A23591
 A:Molecule type: DNA
 A:Residues: 1-26 'EH', 28-33 'TL', 36-54 'R', 56-153 <KEI>
 A:Cross-references: GB:X04118; NID:g21521; PIDD:CAA27730.1; PIDD:g21522
 C:Genetics:
 A:Gene: IIR
 A:Introns: 18/1
 C:Superfamily: potato proteinase inhibitor PTI
 C:Keywords: serine proteinase inhibitor
 P1-25/Domin: signal sequence #status predicted <SIG>
 F15-106/Product: proteinase inhibitor PCI-I #status experimental <MAT>

Query Match 65.0%; Score 78; DB 1; Length 153;
 Best Local Similarity 59.1%; Pred. No. 0.0009; Indels 0; Gaps 0;
 Matches 13; Conservative 1; Mismatches 8;

QY 2 CPXXEKKNDRICTNCCAGXKG 23
 DB 45 CPRSEGNMRICITNCCAGXKG 66

RESULT 5

T07597
 C:Species: Solanum tuberosum (potato)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
 C:Accession: T07597
 R:Lee, J.S.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S24965
 A:Accession: T07597
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-158 <LEE>
 A:Cross-references: EMBL:Z12753; NID:921553; PIDD:CAA78277.1; PIDD:g21554
 A:Experimental source: cv. Russet Burbank
 C:Genetics:
 A:Introns: 17/1
 A:Superfamily: potato proteinase inhibitor PTI
 C:Keywords: serine proteinase inhibitor

Query Match 65.0%; Score 78; DB 2; Length 158;
 Best Local Similarity 59.1%; Pred. No. 0.00092; Indels 0; Gaps 0;
 Matches 13; Conservative 1; Mismatches 8;

QY 2 CPXXEKKNDRICTNCCAGXKG 23
 DB 45 CPRSEGNMRICITNCCAGXKG 66

RESULT 6

T07011
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
 C:Accession: T07011
 R:Gadea, J.; Mayda, E.; Conejero, V.; Vera, P.
 Mol. Plant Microbe Interact. 9, 409-415, 1996
 A>Title: Characterization of defense-related genes ectopically expressed in viroid-infected plants
 A:Reference number: Z15859; MUID:96252900; PMID:8672818
 A:Accession: T07011
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-201 <GMD>
 A:Cross-references: EMBL:X94946; NID:g1161571; PIDD:CAA64416.1; PIDD:g1161572
 A:Experimental source: cultivar Rutgers, leaf
 C:Genetics:
 A:Gene: cev157
 A:Function:
 A:Description: involved in plant defensive responses
 C:Superfamily: potato proteinase inhibitor PTI
 C:Keywords: serine proteinase inhibitor
 F1-23/Domin: signal sequence #status predicted <SIG>

F:24-201/Product: proteinase inhibitor II #status predicted <MAT>

Query Match 62.5%; Score 75; DB 2; Length 201;
Best Local Similarity 54.5%; Pred. No. 0.0028;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEEKNDRICTNCCAGXKG 23
DB 42 CPGSEGSPTNCTCCSGYKG 63

RESULT 7

S72492

probable proteinase inhibitor precursor - tomato

N:Alternate names: AT2 protein

C:Species: Lycopersicon esculentum (tomato)

C>Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 07-May-1999

C:Accession: S72492

R:Brandstaedter, J.; Rosabach, C.; Theres, K.

Mol. Gen. Genet. 252, 146-154, 1996

A:Title: Expression of genes for a defensin and a proteinase inhibitor in specific areas

A:Reference number: S72491; MUID:96397493; PMID:8804387

A:Accession: S72492

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-146

A:Experimental source: shoot, cv. Moneymaker

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: serine proteinase inhibitor

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-146/Product: probable proteinase inhibitor #status predicted <MAT>

F:93/Inhibitory site: Arg (trypsin) #status predicted

F:94/Inhibitory site: Arg (trypsin) #status predicted

Query Match 62.1%; Score 74.5; DB 2; Length 146;
Best Local Similarity 59.1%; Pred. No. 0.0026;
Matches 13; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 2 CPXXEEKNDRICTNCCAGXKG 23

DB 105 CP-GNKRSEGRICNCCAGSKG 125

RESULT 8

JQ2269

trypsin inhibitor-1 - common tobacco

N:Alternate names: proteinase inhibitor

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 31-Mar-2000

C:Accession: JQ2269

R:Pearce, G.; Johnson, S.; Ryan, C.A.

Plant Physiol. 102, 639-644, 1993

A:Title: Purification and characterization from tobacco (Nicotiana tabacum) leaves of si

A:Reference number: JQ2269; MUID:94151442; PMID:8108514

A:Accession: JQ2269

A:Molecule type: protein

A:Residues: 1-53 <PEA>

A:Experimental source: leaf

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: serine proteinase inhibitor

Query Match 61.7%; Score 74; DB 2; Length 53;
Best Local Similarity 92.3%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 DRICNCCAGXKG 23

DB 1 DRICNCCAGTKG 13

RESULT 9

S56662

proteinase inhibitor II precursor - common tobacco

N:Alternate names: serine proteinase inhibitor II
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C:Accession: S56662; S56663
R:Balandin, T.; van der Does, C.; Albert, J.M.B.; Bol, J.F.; Linthorst, H.J.M.
Plant Mol. Biol. 27, 1197-1204, 1995
A:Title: Structure and induction pattern of a novel proteinase inhibitor class II gene of
A:Reference number: S56662; MUID:95284369; PMID:7766901
A:Accession: S56662
A:Molecule type: DNA
A:Residues: 1-197 <BAL>
A:Cross-references: EMBL:Z29537; NID:9453973; PIDN:CAA82652.1; PID:9453974
A:Experimental source: strain Sameun NN; leaf
A:Accession: S56663
A:Molecule type: DNA
A:Residues: 25-197 <BAW>
A:Cross-references: EMBL:Z29537
A:Experimental source: strain Sameun NN; leaf
C:Genetics:
A:Gene: pi2-1
A:Introns: 16/1
C:Function:
A:Description: involved in plant defensive responses
A:Note: not expressed in leaves of healthy plants; expression induced in leaves subjected
C:Superfamily: potato proteinase inhibitor PTI
C:Keywords: serine proteinase inhibitor
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-197/Product: proteinase inhibitor II #status predicted <MAT>

Query Match 61.7%; Score 74; DB 2; Length 197;
Best Local Similarity 54.5%; Pred. No. 0.0038;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEEKNDRICTNCCAGXKG 23

DB 39 CPRSQGTPDDPICTCCAGYKG 60

RESULT 10

S24973

proteinase inhibitor II - potato

C:Species: Solanum tuberosum (potato)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999

C:Accession: S24973

R:Choi, Y.; Kim, J.W.; Lee, J.S.

submitted to the EMBL Data Library, July 1992

A:Description: Characterization of a potato proteinase inhibitor II gene that is expressed

A:Reference number: S24973

A:Accession: S24973

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-147 <CHO>

A:Cross-references: EMBL:Z13992; NID:921555; PIDN:CAA78383.1; PID:921556

C:Genetics:

A:Introns: 18/1

C:Superfamily: potato proteinase inhibitor PTI

Query Match 57.5%; Score 69; DB 2; Length 147;
Best Local Similarity 50.0%; Pred. No. 0.015;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXXEEKNDRICTNCCAGXKG 23

DB 40 CPRSEGSPTNCTCCSGYKG 61

RESULT 11

S43105

proteinase inhibitor II - potato

C:Species: Solanum tuberosum (potato)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999

C:Accession: S43105

R:Murray, C.; Christeller, J.T.

submitted to the EMBL Data Library, March 1994

A:Reference number: S43105

A:Accession: S43105

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-154 <MIR>

A:Cross-references: EMBL:X78275; NID:G467609; PIDN:CAAS5082.1; PID:G467610

C:Genetics:

A:Introns: 18/1

C:Superfamily: potato proteinase inhibitor PTI

Query Match

Best Local Similarity

Matches

Score 69; DB 2; Length 154;

Pred. No. 0.016;

Conservative 2; Mismatches 9; Indels 0; Gaps 0;

CPXHEKNDRICTNCCAGXKG 23

DB 46 CPXHEKNDRICTNCCAGXKG 67

RESULT 12

XREPOT

proteinase inhibitor PTI - potato

C:Species: Solanum tuberosum (potato)

C:Date: 06-Jul-1982 #sequence_revision 06-Jul-1982 #text_change 16-Aug-1996

C:Accession: A01318

R:Rass, G.M.; Hermodson, M.A.; Ryan, C.A.; Gentry, L.

Biochemistry 21, 752-756, 1982

A:Title: Primary structures of two low molecular weight proteinase inhibitors from potato

A:Reference number: A90465; PMID:8218263; PMID:7074039

A:Accession: A01318

A:Molecule type: protein

A:Residues: 1-51 <HMS>

A>Note: Arg-38 is probably the site of interaction with trypsin

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: serine proteinase inhibitor

Query Match

Best Local Similarity

Matches

Score 67; DB 1; Length 51;

Pred. No. 0.013;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 RICTNCCAGXKG 23

DB 1 RICTNCCAGXKG 12

RESULT 13

B24048

proteinase inhibitor II precursor - tomato

N:Alternate names: wound-induced proteinase inhibitor II

C:Species: Lycopersicon esculentum (tomato)

C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 20-Sep-1999

C:Accession: B24048

J:Graham, J.S.; Pearce, G.; Merryweather, J.; Titani, K.; Ericsson, L.; Ryan, C.A.

J. Biol. Chem. 260, 6555-6560, 1985

A:Reference number: A92531; PMID:85207657; PMID:2987227

A:Accession: B24048

A:Molecule type: mRNA

A:Residues: 1-148 <GRA>

C:Comment: The source of this protein was tomato leaves.

C:Comment: Mechanical damage (i.e., insect chewing) to this plant results in the system

inhibitors I and II. Accumulating in the central vacuole of the leaf cells, these potent

C:Comment: This protein is a potent inhibitor of both trypsin and chymotrypsin.

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: serine proteinase inhibitor

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-148/Product: proteinase inhibitor II #status predicted <MAT>

F:26-82/Domain: trypsin-inhibitory <TRY>

F:83-141/Region: duplication

F:83-148/Domain: chymotrypsin-inhibitory <CHY>

F:30/inhibitory site: Arg (trypsin) #status predicted

F:87/inhibitory site: Phe (chymotrypsin) #status predicted

Query Match

Best Local Similarity

Matches

Score 66; DB 2; Length 148;

Pred. No. 0.039;

Conservative 2; Mismatches 9; Indels 0; Gaps 0;

2 CPXHEKNDRICTNCCAGXKG 23

DB 40 CPXHEKNDRICTNCCAGXKG 61

RESULT 14

XREPOT

proteinase inhibitor IIA - potato (fragment)

C:Species: Solanum tuberosum (potato)

C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Aug-1996

C:Accession: A01320

R:Iwasaki, T.; Kiyohara, T.; Yoshikawa, M.

J. Biochem. 79, 381-391, 1976

A:Title: Amino acid sequence of an active fragment of potato proteinase inhibitor IIA.

A:Reference number: A01320; PMID:76190064; PMID:1270410

A:Accession: A01320

A:Molecule type: protein

A:Residues: 1-45 <IWA>

A>Note: this active fragment inhibits trypsin strongly and chymotrypsin temporarily;

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: serine proteinase inhibitor

F:10-24,14-35,20-43/Dipeptide bonds: #status experimental

F:32/inhibitory site: Lys (trypsin) #status experimental

Query Match

Best Local Similarity

Matches

Score 58; DB 1; Length 45;

Pred. No. 0.21;

Conservative 1; Mismatches 6; Indels 0; Gaps 0;

6 EEKNDRICTNCCAGXKG 23

DB 2 EEPENRITNCCAGXKG 19

RESULT 15

TIEOI

proteinase inhibitor - eggplant

C:Species: Solanum melongena (eggplant, aubergine)

C:Date: 30-Nov-1980 #sequence_revision 30-Nov-1980 #text_change 05-Aug-1994

C:Accession: A01317

R:Richardson, M. 322-326, 1979

FEBS Lett. 104, 322-326, 1979

A:Title: The complete amino acid sequence and the trypsin reactive (inhibitory) site

A:Reference number: A01317; PMID:8004153; PMID:477995

A:Accession: A01317

A:Molecule type: protein

A:Residues: 1-52 <RIC>

A>Note: 2-Leu and 6-Cys were also found

C:Superfamily: potato proteinase inhibitor PTI

C:Keywords: pyroglutamic acid

F:1/Modified site: pyroglutamic carboxylic acid (Gln) #status experimental

F:38/inhibitory site: Arg (trypsin) #status experimental

Query Match

Best Local Similarity

Matches

Score 51; DB 1; Length 52;

Pred. No. 2.1;

Conservative 9; Mismatches 2; Indels 0; Gaps 0;

12 RICTNCCAGXKG 23

DB 1 QICTNNCCARXKG 12

Search completed: December 17, 2002, 10:24:36

Job time : 18 secs

Result No.	Score	Query #		DB	ID	Description
		Match	Length			
1	85	70.8	154	1	IP25_SOLTU	Q41488 solanum tub
2	85	70.8	204	1	IP22_CAPAN	Q49146 capsicum an
3	82	68.3	223	1	IP22_LYCES	Q43170 lycopersico
4	78	65.0	153	1	IP2K_SOLTU	P01080 solanum tub
5	78	65.0	158	1	IP2K_SOLTU	Q00782 solanum tub
6	75	62.5	201	1	IP23_LYCES	Q43502 lycopersico
7	74	61.7	197	1	IP21_TOBAC	Q40561 nicotiana t
8	69	57.5	147	1	IP2T_SOLTU	Q41435 solanum tub
9	69	57.5	147	1	IP2T_SOLTU	Q41435 solanum tub
10	69	57.5	154	1	IP27_SOLTU	Q41489 solanum tub
11	67	55.8	51	1	IP21_SOLTU	Q43652 solanum tub
12	67	55.8	55	1	IP21_CAPAN	P01079 solanum tub
13	66	55.0	148	1	IP21_LYCES	P56615 capsicum an
14	63	52.5	52	1	IPR_SOLME	P05119 lycopersico
15	58	48.3	45	1	IP2A_SOLTU	P01078 solanum mel
16	50	41.7	40	1	IP2B_SOLTU	P01081 solanum tub
17	49	40.8	180	1	TR22_MOUSE	P01082 solanum tub
18	49	40.8	919	1	YB03_YEAST	Q9er62 mus musculu
19	47	39.2	176	1	TR23_MOUSE	P38073 saccharomyc
20	47	39.2	762	1	PDE1_ARATH	Q9er63 mus musculu
21	46	38.3	763	1	GLH1_CABEL	Q9C988 arabidopsis
22	46	38.3	1127	1	TF1G_HUMAN	P34589 caenorhabdi
23	46	38.3	1906	1	KWLS_CHICK	Q9upn9 homo sapien
24	45	37.5	761	1	TOP1_METJJA	P11799 gallus gall
25	44.5	37.1	1609	1	LMG1_HUMAN	Q59046 methanococc
26	44	36.7	760	1	AHM4_ARATH	P11047 homo sapien
27	44	36.7	863	1	LDVR_CHICK	Q9szw5 arabidopsis
28	44	36.7	873	1	LDVR_HUMAN	P98165 gallus gall
29	44	36.7	873	1	LDVR_MOUSE	P98155 homo sapien
30	44	36.7	873	1	LDVR_RABIT	P98156 mus musculu
31	44	36.7	873	1	LDVR_RAT	P35953 oryctolagus
32	44	36.7	1808	1	TENA_CHICK	P98166 rattus norv
33	43.5	36.2	62	1	MT2_CABEL	P17512 caenorhabdi

RESULT 2
IP22_CAPAN STANDARD; PRT; 204 AA.
ID IP22_CAPAN
AC 049146;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Wound-induced proteinase inhibitor II precursor.
GN PIN2
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hot pepper; TISSUE=pericarp;
RA Kim S.-H., Choi D.-S., Lee K.-W.,
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF039398; AAB94771.1; -.
DR HSSP; P01080; 4SG8.
DR InterPro; IPR003465; Prot_inhib.
DR Pfam; PF02428; Prot_inhib_II; 4.
KM Serine protease inhibitor; Repeat; Signal.
FT SIGNAL 1 25
FT CHAIN 26 204 WOUND-INDUCED PROTEINASE INHIBITOR II.
FT REPEAT 25 67 POTENTIAL.
FT REPEAT 68 125
FT REPEAT 126 183
FT REPEAT 184 204 4 (PARTIAL).
SQ SEQUENCE 204 AA; 22187 MW; EA9416689E39E05 CRC64;
Query Match 70.8%; Score 85; DB 1; Length 204;
Best Local Similarity 63.6%; Pred. No. 2.6e-05;
Matches 14; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
OY 2 CPXXEKKNDRICTNCCAGXKG 23
DB 84 CPBSEGAENRITCNCCAGRKG 105
RESULT 3
IP22_LYCES STANDARD; PRT; 223 AA.
ID IP22_LYCES
AC 043710;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Proteinase inhibitor type II TR8 precursor.
GN ARPI
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VFN8; TISSUE=seedling root;
RA MEDLINE=94083553; PubMed=7903168;
RT Taylor B.H., Young R.J., Scheuring C.F.;
RT "Induction of a proteinase inhibitor II-class gene by auxin in tomato
roots.";
RT Plant Mol. Biol. 23:1005-1014(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VFN8;
RA MEDLINE=94211920; PubMed=8159801;
RT Young R.J., Scheuring C.F., Harris-Haller L., Taylor B.H.;
RT "An auxin-inducible proteinase inhibitor gene from tomato."
RT Plant Physiol. 104:811-812(1994).
CC -1- INDICATION: BY AUXIN.
CC -1- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L21194; AAA16881.1; -.
DR EMBL; L25128; AAC37397.1; -.
DR HSSP; P01080; 4SG8.
DR InterPro; IPR003465; Prot_inhib.
DR Pfam; PF02428; Prot_inhib_II; 3.
KM Serine protease inhibitor; Repeat; Signal.
FT SIGNAL 1 24
FT CHAIN 25 223 POTENTIAL.
FT REPEAT 24 81
FT REPEAT 86 145
FT REPEAT 152 209
FT ACT_SITE 29 30 INTERACTION WITH TRYPSIN (PROBABLE).
FT ACT_SITE 93 94 INTERACTION WITH TRYPSIN (PROBABLE).
FT ACT_SITE 157 158 INTERACTION WITH TRYPSIN (PROBABLE).
SQ SEQUENCE 223 AA; 24697 MW; 8F6173C4BE36F9E CRC64;
Query Match 68.3%; Score 82; DB 1; Length 223;
Best Local Similarity 59.1%; Pred. No. 7.4e-05;
Matches 13; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
OY 2 CPXXEKKNDRICTNCCAGXKG 23
DB 104 CPLETKRVEGLCTNCCAGXKG 125
RESULT 4
IP2K_SOLTU STANDARD; PRT; 153 AA.
ID IP2K_SOLTU
AC P01080; P11429;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteinase inhibitor type II K precursor (IIK) (PCT-I).
GN PIN2K.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RA Thornburg R.W., An G., Cleveland T.E., Johnson R., Ryan C.A.;
RT "Wound-inducible expression of a potato inhibitor II-chloramphenicol
RT acetyltransferase gene fusion in transgenic tobacco plants."
RT Proc. Natl. Acad. Sci. U.S.A. 84:744-746(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86286579; PubMed=3016659;
RA Keil M., Sanchez-Serrano J., Scheil J., Willmitzer L.;
RT "Primary structure of a proteinase inhibitor II gene from potato
RT (Solanum tuberosum)."
RT Nucleic Acids Res. 14:5641-5650(1986).

```

RN  [3]
RP  SEQUENCE OF 55-106 FROM N.A.
RX  MEDLINE=82182863; PubMed=7074039;
RA  Hass G.M., Hermodson M.A., Ryan C.A., Gentry L.;
RT  "Primary structures of two low molecular weight proteinase inhibitors
RL  from potatoes";
RN  Biochemistry 21:752-756(1982).
RN  [4]
RX  X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 55-106.
RX  MEDLINE=89178636; PubMed=2494344;
RA  Greenblatt H.M., Ryan C.A., James M.N.G.;
RT  "Structure of the complex of Streptomyces griseus proteinase B and
RT  polypeptide chymotrypsin inhibitor-1 from Russet Burbank potato
RT  tubers at 2.1-A resolution.";
RL  J. Mol. Biol. 205:201-228(1989).
CC  -|- FUNCTION: INHIBITOR OF TRYPSIN AND CHYMOTRYPSIN.
CC  -|- INDUCTION: BY WOUNDING.
CC  -|- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M15186; AAA33815.1; -
DR  EMBL; X03778; CAA27408.1; -
DR  EMBL; X04118; CAA27730.1; -
DR* PIR; A23591; A23591.
DR  PIR; A26584; A26584.
DR  PIR; A01319; XKPOCI.
DR  PDB; 4SGB; 15-JUL-90.
DR  InterPro; IPR003465; Prot inhib.
DR  Pfam; PF02428; Prot inhib II; 2.
KW  Serine protease inhibitor; Repeat; Signal; 3D-structure.
FT  SIGNAL 1 25
FT  CHAIN 26 153
FT  CHAIN 55 106
FT  DISULFID 57 94
FT  DISULFID 60 78
FT  DISULFID 61 90
FT  DISULFID 67 103
FT  ACT_SITE 35 36
FT  ACT_SITE 92 93
FT  REPEAT 30 86
FT  REPEAT 87 146
FT  CONFLICT 27 27
FT  CONFLICT 34 35
FT  CONFLICT 55 55
FT  TURN 60 62
FT  STRAND 64 64
FT  TURN 65 66
FT  STRAND 68 70
FT  TURN 72 73
FT  STRAND 75 79
FT  STRAND 82 82
FT  TURN 84 85
FT  STRAND 90 91
FT  STRAND 94 94
FT  TURN 96 97
FT  STRAND 100 102
SQ  SEQUENCE 153 AA; 16505 MW; 8027499652E8BA06 CRC64;

Query Match 65.0%; Score 78; DB 1; Length 153;
Best Local Similarity 59.1%; Pred. No. 0.0002;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEEKNDRICTNCCAGXKG 23
|||:|||||
Db 45 CPRSEGSPEPNTCCAGYKG 66

RN  [3]
RP  SEQUENCE OF 55-106 FROM N.A.
RX  MEDLINE=82182863; PubMed=7074039;
RA  Hass G.M., Hermodson M.A., Ryan C.A., Gentry L.;
RT  "Primary structures of two low molecular weight proteinase inhibitors
RL  from potatoes";
RN  Biochemistry 21:752-756(1982).
RN  [4]
RX  X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 55-106.
RX  MEDLINE=89178636; PubMed=2494344;
RA  Greenblatt H.M., Ryan C.A., James M.N.G.;
RT  "Structure of the complex of Streptomyces griseus proteinase B and
RT  polypeptide chymotrypsin inhibitor-1 from Russet Burbank potato
RT  tubers at 2.1-A resolution.";
RL  J. Mol. Biol. 205:201-228(1989).
CC  -|- FUNCTION: INHIBITOR OF TRYPSIN AND CHYMOTRYPSIN.
CC  -|- INDUCTION: BY WOUNDING.
CC  -|- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M15186; AAA33815.1; -
DR  EMBL; X03778; CAA27408.1; -
DR  EMBL; X04118; CAA27730.1; -
DR* PIR; A23591; A23591.
DR  PIR; A26584; A26584.
DR  PIR; A01319; XKPOCI.
DR  PDB; 4SGB; 15-JUL-90.
DR  InterPro; IPR003465; Prot inhib.
DR  Pfam; PF02428; Prot inhib II; 2.
KW  Serine protease inhibitor; Repeat; Signal; 3D-structure.
FT  SIGNAL 1 25
FT  CHAIN 26 153
FT  CHAIN 55 106
FT  DISULFID 57 94
FT  DISULFID 60 78
FT  DISULFID 61 90
FT  DISULFID 67 103
FT  ACT_SITE 35 36
FT  ACT_SITE 92 93
FT  REPEAT 30 86
FT  REPEAT 87 146
FT  CONFLICT 27 27
FT  CONFLICT 34 35
FT  CONFLICT 55 55
FT  TURN 60 62
FT  STRAND 64 64
FT  TURN 65 66
FT  STRAND 68 70
FT  TURN 72 73
FT  STRAND 75 79
FT  STRAND 82 82
FT  TURN 84 85
FT  STRAND 90 91
FT  STRAND 94 94
FT  TURN 96 97
FT  STRAND 100 102
SQ  SEQUENCE 153 AA; 16505 MW; 8027499652E8BA06 CRC64;

Query Match 65.0%; Score 78; DB 1; Length 153;
Best Local Similarity 59.1%; Pred. No. 0.0002;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEEKNDRICTNCCAGXKG 23
|||:|||||
Db 45 CPRSEGSPEPNTCCAGYKG 66

```

```

RESULT 5
IP2X_SOLTU
ID IP2X_SOLTU STANDARD; PRT; 158 AA.
AC Q00782;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Proteinase inhibitor type II precursor.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4113;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Russet Burbank;
RA Choi Y., Moon Y., Lee J.S.;
RT "Primary structure of two proteinase inhibitor II genes closely linked
RL in the potato genome.";
RL Korean J. Biochem. 23:214-220(1990).
CC -|- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; Z12753; CAA78277.1; -
DR  HSSP; P01080; 4SGB.
DR  InterPro; IPR003465; Prot inhib.
DR  Pfam; PF02428; Prot inhib II; 3.
KW  Serine protease inhibitor; Repeat; Signal.
FT  SIGNAL 1 24
FT  CHAIN 25 158
FT  DISULFID 57 94
FT  DISULFID 60 78
FT  DISULFID 61 90
FT  DISULFID 67 103
FT  ACT_SITE 35 36
FT  ACT_SITE 92 93
FT  REPEAT 29 86
FT  REPEAT 87 146
FT  SEQUENCE 158 AA; 17131 MW; C84DF44B015F248E CRC64;

Query Match 65.0%; Score 78; DB 1; Length 158;
Best Local Similarity 59.1%; Pred. No. 0.00021;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEEKNDRICTNCCAGXKG 23
|||:|||||
Db 45 CPRSEGSPEPNTCCAGYKG 66

RESULT 6
IP23_LYCES
ID IP23_LYCES STANDARD; PRT; 201 AA.
AC Q43502;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Proteinase inhibitor type II CEVI57 precursor.
GN CEVI57.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Rutgers; TISSUE=leaf;
RX MEDLINE=96252900; PubMed=9672818;
RA Gadea J., Mayda E., Conejero V., Vera P.;
RT "Characterization of defense-related genes ectopically expressed in
RL viroid-infected tomato plants."
RM Mol. Plant Microbe Interact. 9:409-415(1996).
CC -I- INDUCTION: BY VIROID INFECTION.
CC -I- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X94946; CAA64416.1; -.
DR HSSP; P01080; 4SGB.
DR InterPro; IPR003465; Prot_inhb.
DR Pfam; PF02428; Prot_inhb_II; 3.
DR Serine protease inhibitor; Repeat; Signal.
KW CHAIN 1 23
FT SIGNAL 1 23
FT CHAIN 1 23
FT REPEAT 27 83
FT REPEAT 84 143
FT REPEAT 144 199
FT ACT_SITE 32 33
FT ACT_SITE 147 148
FT ACT_SITE 201 AA; 21419 MW; A3FCAB937373D8590 CRC64;
SQ SEQUENCE 201 AA; 21419 MW; A3FCAB937373D8590 CRC64;

Query Match 62.5%; Score 75; DB 1; Length 201;
Best local Similarity 54.5%; Pred. No. 0.00066;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCCAGKKG 23
DB 42 CPSESGSPENPITCCSGYKG 63

RESULT 7
IP21_TOBAC STANDARD; PRT; 197 AA.
AC Q40561;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Proteinase inhibitor type II precursor.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Assteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
ON NCBI_TaxID=4097;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Samsun NN; TISSUE=leaf;
RX MEDLINE=95284369; PubMed=7766901;
RA Balandin M.T., van der Does C., Albert J.M., Bol J.F.,
RA Linthorst H.J.M.;
RT "Structure and induction pattern of a novel proteinase inhibitor
RT class II gene of tobacco."
RL Plant Mol. Biol. 27:1197-1204(1995).
CC -I- INDUCTION: LOCALLY INDUCED IN LEAVES SUBJECTED TO DIFFERENT TYPES
CC OF STRESS (TMV INFECTION, WOUNDING, UV IRRADIATION).
CC -I- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29537; CAA82652.1; -.
DR HSSP; P01080; 4SGB.
DR InterPro; IPR003465; Prot_inhb.
DR Pfam; PF02428; Prot_inhb_II; 3.
DR Serine protease inhibitor; Repeat; Signal.
KW CHAIN 1 24
FT SIGNAL 1 24
FT CHAIN 1 24
FT REPEAT 25 197
FT REPEAT 80 24
FT REPEAT 81 140
FT REPEAT 141 196
FT ACT_SITE 29 30
FT ACT_SITE 197 AA; 20984 MW; F9CD5C84267A9710 CRC64;
SQ SEQUENCE 197 AA; 20984 MW; F9CD5C84267A9710 CRC64;

Query Match 61.7%; Score 74; DB 1; Length 197;
Best local Similarity 54.5%; Pred. No. 0.0009;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCCAGKKG 23
DB 39 CPNSQTPDDPITCCAGYKG 60

RESULT 8
IP21_SOLITU STANDARD; PRT; 147 AA.
AC Q41435;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteinase inhibitor type II T precursor.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Assteridae; euasterids I; Solanales; Solanaceae; Solanum.
ON NCBI_TaxID=4113;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Russet Burbank;
RA Park S., Thornburg R.W.;
RT "Isolation and characterization of a proteinase inhibitor II gene that
RT is not wound-inducible."
RL (In) Plant Gene Register FGR96-007.
CC -I- FUNCTION: INHIBITOR OF TRYPSIN AND CHYMOTRYPSIN.
CC -I- INDUCTION: NOT INDUCED BY WOUNDING.
CC -I- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U45450; AAD09849.1; -.
DR HSSP; P01080; 4SGB.
DR InterPro; IPR003465; Prot_inhb.
DR Pfam; PF02428; Prot_inhb_II; 2.
DR Serine protease inhibitor; Repeat; Signal.
KW CHAIN 1 25
FT SIGNAL 1 25
FT CHAIN 1 25
FT DISULFID 52 89
FT DISULFID 55 73
FT DISULFID 56 85
FT DISULFID 62 98
FT ACT_SITE 30 31
FT ACT_SITE 31

```

```

FT ACT SITE 87 88 INTERACTION WITH CHYMOTRYPSIN (PROBABLE).
FT REPEAT 25 82
FT REPEAT 83 142
SQ SEQUENCE 147 AA; 15860 MW; 895D2F3102B8C2AE CRC64;

Query Match 57.5%; Score 69; DB 1; Length 147;
Best Local Similarity 50.0%; Pred. No. 0.0036;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXEEKKNDRICTNCCAGXKG 23
DB 40 CPRSEGSTNPICNCCSGYKG 61

RESULT 9
IP2Y_SOLU
ID IP2Y_SOLU STANDARD; PRT; 147 AA.
AC Q41489;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Proteinase inhibitor type II precursor.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Russet Burbank;
RA Choi Y., Kim J.W., Lee J.S.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
CC -| SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Z13992; CAA78383.1; --
CC HSP; P01080; 4SGB.
CC InterPro; IPR003465; Prot. inhib.
CC Pfam; PF02428; Prot. inhib. II; 2.
CC Serine protease inhibitor; Repeat; Signal.
CC SIGNAL 1 25 POTENTIAL.
CC CHAIN 26 147 PROTEINASE INHIBITOR TYPE II.
CC DISULFID 52 89 BY SIMILARITY.
CC DISULFID 55 73 BY SIMILARITY.
CC DISULFID 56 85 BY SIMILARITY.
CC DISULFID 62 98 BY SIMILARITY.
CC ACT_SITE 30 31 INTERACTION WITH TRYPSIN (PROBABLE).
CC ACT_SITE 87 88 INTERACTION WITH CHYMOTRYPSIN (PROBABLE).
CC REPEAT 25 81 1.
CC REPEAT 82 141 2.
CC SEQUENCE 147 AA; 15936 MW; 703456551B54F968 CRC64;

Query Match 57.5%; Score 69; DB 1; Length 147;
Best Local Similarity 50.0%; Pred. No. 0.0036;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXEEKKNDRICTNCCAGXKG 23
DB 40 CPRSEGSTNPICNCCSGYKG 61

RESULT 10
IP27_SOLU
ID IP27_SOLU STANDARD; PRT; 154 AA.
AC Q43652;

```

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Proteinase inhibitor type II CM7 precursor.
GN PIN2-CM7.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Arran Banner; TISSUE=Leaf;
RX MEDLINE=95148744; PubMed=7846166;
RA Murray C., Christeller J.T.;
RL "Genomic nucleotide sequence of a proteinase inhibitor II gene.";
RL Plant Physiol. 106:1681-1691 (1994).
CC -| SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X78275; CAA55082.1; --
CC HSP; P01080; 4SGB.
CC InterPro; IPR003465; Prot. inhib.
CC Pfam; PF02428; Prot. inhib. II; 2.
CC Serine protease inhibitor; Repeat; Signal.
CC SIGNAL 1 31 POTENTIAL.
CC CHAIN 32 154 PROTEINASE INHIBITOR TYPE II CM7.
CC DISULFID 58 95 BY SIMILARITY.
CC DISULFID 61 79 BY SIMILARITY.
CC DISULFID 62 91 BY SIMILARITY.
CC DISULFID 68 104 BY SIMILARITY.
CC ACT_SITE 36 37 INTERACTION WITH TRYPSIN (PROBABLE).
CC ACT_SITE 93 94 INTERACTION WITH CHYMOTRYPSIN (PROBABLE).
CC REPEAT 31 87 1.
CC REPEAT 88 147 2.
CC SEQUENCE 154 AA; 16868 MW; 9BE8CCB7A26099C2 CRC64;

Query Match 57.5%; Score 69; DB 1; Length 154;
Best Local Similarity 50.0%; Pred. No. 0.0038;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 CPXEEKKNDRICTNCCAGXKG 23
DB 46 CPRSEGSTNPICNCCSGYKG 67

RESULT 11
IP21_SOLU
ID IP21_SOLU STANDARD; PRT; 51 AA.
AC P01079;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Proteinase inhibitor PTI.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX MEDLINE=82182863; PubMed=7074039;
RA Haas G.M., Hermodson M.A., Ryan C.A., Gentry L.;
RT "Primary structures of two low molecular weight proteinase inhibitors
RT from potatoes."

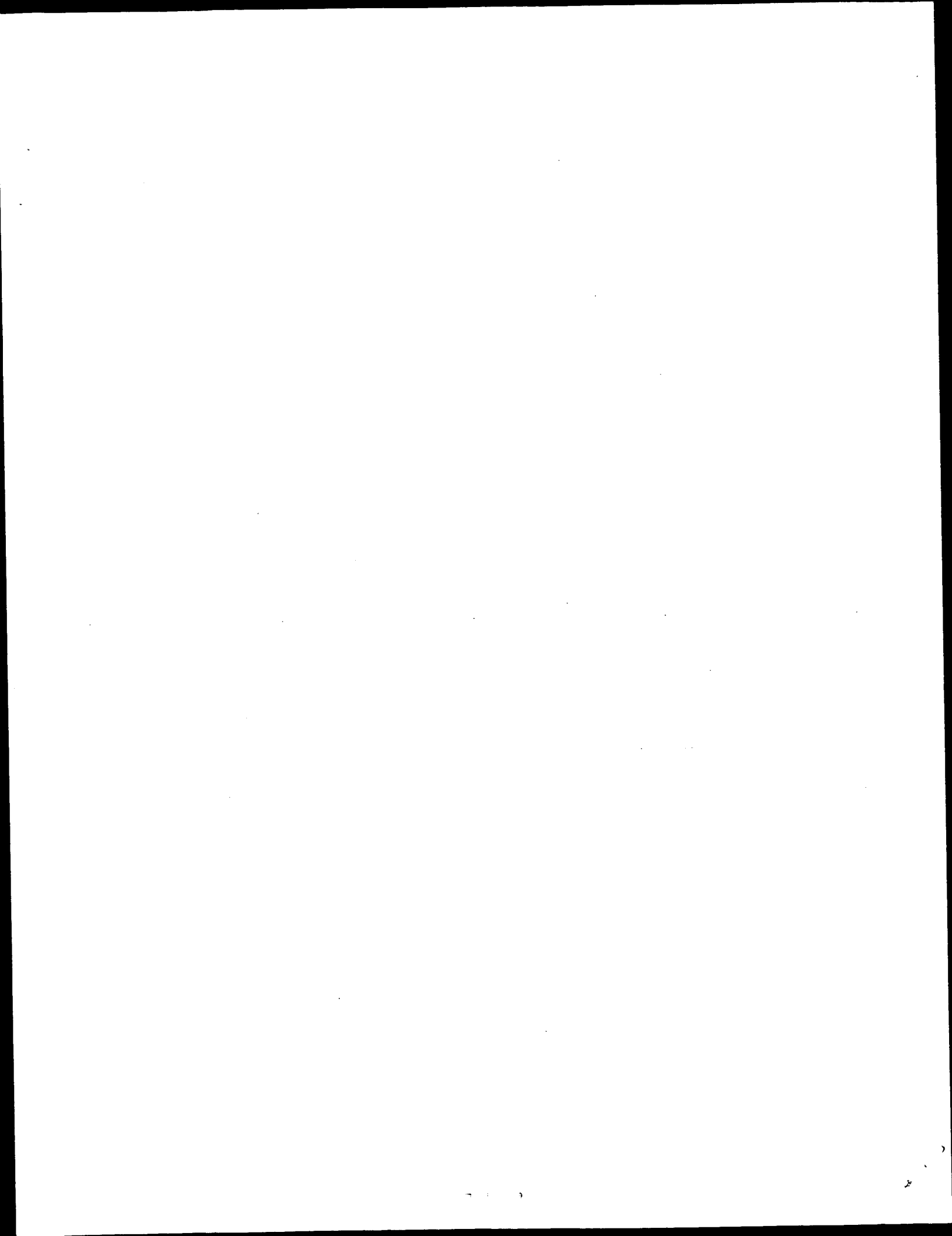
```

RL Biochemistry 21:752-756(1982).
 CC -1- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
 CC FAMILY.
 DR PIR: A01318; XKPOT.
 DR HSSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot_inhb.
 DR Pfam; PF02428; Prot_inhb_II; 2.
 KW Serine protease inhibitor.
 FT DISULFID 3 40 BY SIMILARITY.
 FT DISULFID 6 24 BY SIMILARITY.
 FT DISULFID 7 36 BY SIMILARITY.
 FT DISULFID 13 49 BY SIMILARITY.
 FT ACT_SITE 38 39 INTERACTION WITH TRYPSIN (PROBABLE).
 SQ SEQUENCE 51 AA; 5602 MW; C8491BF60CD4075 CRC64;
 QY Query Match Score 67; DB 1; Length 51;
 Best Local Similarity 91.7%; Pred. No. 0.003;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 12 RICTNCCAGXKG 23
 Db 1 RICTNCCAGXKG 12
 RESULT 12
 ID IP21 CAPAN STANDARD; PRT; 55 AA.
 AC P56615;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Proteinase inhibitor PSI-1.1.
 OS Capsicum annuum (Bell pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
 OC NCBI_TaxID=4072;
 OX [1]
 RN SEQUENCE.
 RP TISSUE=Seed;
 RC MEDLINE=97104335; PubMed=8948493;
 RX Antcheva N., Patchy A., Athanasiadis A., Tchobanov B., Zakhariev S.,
 RA Pongor S.;
 RT "Primary structure and specificity of a serine proteinase inhibitor
 RT from paprika (Capsicum annuum) seeds.";
 RL Biochim. Biophys. Acta 1298:95-101(1996).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RA Antcheva N., Patchy A., Athanasiadis A., Tchobanov B., Zakhariev S.,
 RA Pongor S.;
 RT "Isolation and characterization of a major serine proteinase inhibitor
 RT from paprika (Capsicum annuum) seeds.";
 RL (in) Ramage R., Epton R. (eds.);
 RL Peptides 1996, pp.209-210, Mayflower Scientific, Kingswinford (1996).
 CC -1- FUNCTION: POTENT INHIBITOR OF TRYPSIN AND A WEAKER INHIBITOR OF
 CC CHYMOTRYPSIN AND PRONASE E.
 CC -1- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
 CC FAMILY.
 CC HSSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot_inhb.
 DR Pfam; PF02428; Prot_inhb_II; 2.
 KW Serine protease inhibitor.
 FT DISULFID 3 40
 FT DISULFID 6 24
 FT DISULFID 7 36
 FT DISULFID 13 49
 FT ACT_SITE 38 39
 SQ SEQUENCE 55 AA; 6061 MW; E401FF2DB8F3CE08 CRC64;
 QY Query Match Score 67; DB 1; Length 55;
 Best Local Similarity 91.7%; Pred. No. 0.0032;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 12 RICTNCCAGXKG 23
 Db 1 RICTNCCAGXKG 12
 RESULT 13
 ID IP21 LYCES STANDARD; PRT; 148 AA.
 AC P05119;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Wound-induced proteinase inhibitor II precursor.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Leaf;
 RC MEDLINE=85207658; PubMed=3838986;
 RX Graham U.S., Pearce G., Merryweather J., Titani K., Ericsson L.H.,
 RA Ryan C.A.;
 RT "Wound-induced proteinase inhibitors from tomato leaves. II. The
 RT cDNA-decoded primary structure of pre-inhibitor II.";
 RL J. Biol. Chem. 260:6561-6564(1985).
 CC -1- FUNCTION: POTENT INHIBITOR OF BOTH TRYPSIN AND CHYMOTRYPSIN.
 CC -1- INDUCTION: MECHANICAL DAMAGE (I.E., INSECT CHEWING) TO THIS PLANT
 CC RESULTS IN THE SYSTEMIC RELEASE OF A FACTOR FROM THE WOUND SITE.
 CC WITHIN THE LEAVES IT INDUCES THE CYTOPLASMIC SYNTHESIS OF
 CC PROTEINASE INHIBITORS I AND II.
 CC -1- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K03291; AAA34201.1; -
 DR PIR; B24048; B24048.
 DR HSSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot_inhb.
 DR Pfam; PF02428; Prot_inhb_II; 2.
 KW Serine protease inhibitor; Repeat; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 148
 FT REPEAT 26 81
 FT REPEAT 83 141
 FT ACT_SITE 30 31
 FT ACT_SITE 87 88
 SQ SEQUENCE 148 AA; 16293 MW; 468A2F653971AANC CRC64;
 QY Query Match Score 66; DB 1; Length 148;
 Best Local Similarity 50.0%; Pred. No. 0.0096;
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 OY 2 CPXAEKNDRICTNCCAGXKG 23
 Db 40 CPXAEKNDRICTNCCAGXKG 61
 RESULT 14
 ID IPR SOLME STANDARD; PRT; 52 AA.
 AC P01078;
 DT 21-JUL-1986 (Rel. 01, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Proteinase inhibitor.
 OS Solanum melongena (Eggplant) (Aubergine).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 ON NCBI_TaxID=4111;
 RX MEDLINE=80004163; PubMed=477995;
 RA Richardson M.;
 RT "The complete amino acid sequence and the trypsin reactive
 RT (inhibitory) site of the major proteinase inhibitor from the fruits
 RT of aubergine (Solanum melongena L.).";
 RL FEBS Lett. 104:322-326(1979).
 CC -|- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
 CC FAMILY.
 DR PIR; A01317; TIE01.
 DR HSP; P01080; 4SGH.
 DR InterPro; IPR001230; Prenyl site.
 DR InterPro; IPR003465; Prot_inhib.
 DR Pfam; PF02428; Prot_inhib_II; 2.
 KW Serine protease inhibitor.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 3 40 BY SIMILARITY.
 FT DISULFID 6 24 BY SIMILARITY.
 FT DISULFID 7 36 BY SIMILARITY.
 FT DISULFID 13 49 BY SIMILARITY.
 FT ACT_SITE 38 39 INTERACTION WITH TRYPSIN.
 FT VARIANT 2 2 I -> L.
 FT VARIANT 6 6 C -> N.
 SQ SEQUENCE 52 AA; 5579 MW; C5B8C2F1D9B9217 CRC64;
 Query Match 52.5%; Score 63; DB 1; Length 52;
 Best Local Similarity 83.3%; Pred. No. 0.011;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 12 RICTNCCAGXKG 23
 Db 1 QICTNCCAGXKG 12
 RESULT 15
 IP2A_SOLTU STANDARD; PRT; 45 AA.
 AC P01081;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Proteinase inhibitor IIA (Fragment).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 ON NCBI_TaxID=4113;
 RX MEDLINE=76190064; PubMed=1270410;
 RA Iwasaki T., Kiyohara T., Yoshikawa M.;
 RT "Amino acid sequence of an active fragment of potato proteinase
 RT inhibitor Iia.";
 RL J. Biochem. 79:381-391(1976).
 CC -|- FUNCTION: INHIBITS TRYPSIN STRONGLY AND CHYMOTRYPSIN TEMPORARILY.
 CC -|- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
 CC FAMILY.
 DR PIR; A01320; XKP02A.
 DR HSP; P01080; 4SGH.
 DR InterPro; IPR003465; Prot_inhib.
 DR Pfam; PF02428; Prot_inhib_II; 1.
 KW Serine protease inhibitor.
 FT NON_TER 1 1
 FT DISULFID 10 24

FT DISULFID 14 35
 FT DISULFID 20 43
 FT ACT_SITE 32 33
 FT NON_TER 45 45
 SQ SEQUENCE 45 AA; 4955 MW; E0FD71BF41499F98 CRC64;
 Query Match 48.3%; Score 58; DB 1; Length 45;
 Best Local Similarity 61.1%; Pred. No. 0.05;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 Oy 6 BEKNDRICTNCCAGXKG 23
 Db 2 EGSPENRICTNCCAGXKG 19
 Search completed: December 17, 2002, 10:23:37
 Job time : 12 secs



GenCore version 5.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: December 17, 2002, 10:22:03 ; Search time 29 seconds
 (without alignments)
 163.417 Million cell updates/sec

Title: US-09-812-502A-16
 Perfect score: 120
 Sequence: 1 XCPXXEKKNDRICTNCCAGXKG 23

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
 Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : SPTREMBL 21.*
 1: sp archaea.*
 2: sp bacteria.*
 3: sp fungi.*
 4: sp human.*
 5: sp invertebrate.*
 6: sp mammal.*
 7: sp mhc.*
 8: sp organelle.*
 9: sp phage.*
 10: sp plant.*
 11: sp rodent.*
 12: sp virus.*
 13: sp vertebrate.*
 14: sp unclassified.*
 15: sp rvirus.*
 16: sp bacteriaph.*
 17: sp archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	95.8	281	10 Q9SQ77	Q9sq77 nicotiana a
2	115	95.8	397	10 Q40378	Q40378 nicotiana a
3	109	90.8	390	10 Q9SDW7	Q9sdw7 nicotiana g
4	109	90.8	506	10 Q9SDW8	Q9sdw8 nicotiana g
5	85	70.8	204	10 Q9SDLA	Q9sdla capsicum an
6	78	65.0	126	10 Q92735	Q92735 solanum tub
7	75	62.5	109	10 Q9W2K1	Q9w2k1 solanum ame
8	73	60.8	148	10 Q9W2K6	Q9w2k6 capsicum an
9	71	59.2	143	10 Q9W514	Q9w514 capsicum an
10	70	58.3	52	10 P83241	P83241 capsicum an
11	54	45.0	46	10 Q94192	Q94192 atropa bell
12	53	44.2	417	5 Q9W564	Q9w564 drosophila
13	52	43.3	361	10 Q9FG64	Q9fg64 arabidopsis
14	52	43.3	377	16 Q9AK62	Q9ak62 streptomyces
15	51	42.5	168	10 Q9LNG2	Q9lng2 arabidopsis
16	51	42.5	303	10 Q9C727	Q9c727 arabidopsis

17	51	42.5	503	5 Q9U018	Q9u018 giardia lam
18	50.5	42.1	45	10 Q9LW92	Q9lw92 nicotiana t
19	49	40.8	3443	12 Q11979	Q11979 maize chlor
20	48	40.0	775	5 Q8T314	Q8t314 paramesium
21	47.5	39.6	418	5 Q26662	Q26662 strongyloce
22	47	39.2	59	12 Q8UZC8	Q8uzc8 cercopithec
23	47	39.2	132	13 Q98SM2	Q98sm2 rana catesb
24	47	39.2	490	11 Q8VCL9	Q8vcl9 mus musculu
25	47	39.2	491	4 Q9H586	Q9h586 homo sapien
26	47	39.2	491	4 Q8TAJ4	Q8taj4 homo sapien
27	47	39.2	491	11 Q9JIY2	Q9jiy2 mus musculu
28	47	39.2	550	5 Q9N457	Q9n457 caenorhabdi
29	47	39.2	869	13 Q42126	Q42126 xenopus lae
30	47	39.2	974	5 Q966L9	Q966l9 caenorhabdi
31	47	39.2	974	5 Q27376	Q27376 caenorhabdi
32	46	38.3	579	13 Q9PTT7	Q9ptt7 brachydanio
33	46	38.3	579	13 Q9YI00	Q9yi00 brachydanio
34	46	38.3	579	13 Q9W6E4	Q9w6e4 brachydanio
35	46	38.3	763	5 Q22873	Q22873 caenorhabdi
36	46	38.3	1074	5 Q964D1	Q964d1 entamoeba h
37	45	37.5	132	13 Q98SM1	Q98sm1 rana catesb
38	45	37.5	164	10 Q9MA78	Q9ma78 arabidopsis
39	45	37.5	383	16 Q9K5N2	Q9k5n2 bacillus ha
40	45	37.5	650	13 Q9I9A8	Q9i9a8 xenopus lae
41	45	37.5	857	15 Q66956	Q66956 feline immu
42	45	37.5	873	16 Q98LX0	Q98lx0 thizobium l
43	44	36.7	77	13 Q96233	Q96233 xenopus lae
44	44	36.7	92	16 Q9X1L9	Q9x1l9 thermotoga
45	44	36.7	101	5 Q9U341	Q9u341 caenorhabdi

ALIGNMENTS

RESULT 1
 ID Q9SQ77 PRELIMINARY; PRT; 281 AA.
 AC Q9SQ77;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Proteinase inhibitor.
 OS Nicotiana glauca (Winged tobacco) (persian tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4087;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=STIGMA;
 RX MEDLINE=20252525; PubMed=10794532;
 RA Miller E.A., Lee M.C.S., Atkinson A.H.O., Anderson M.A.;
 RT "Identification of a novel four-domain member of the proteinase
 RT inhibitor II family from the stigmas of Nicotiana glauca";
 RL Plant Mol. Biol. 42:329-333(2000).
 DR EMBL; AF105340; AAF14181.1; -
 DR HSSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot. inhib.
 DR Pfam; PF02428; Prot. inhib-II; 4.
 SQ SEQUENCE 281 AA; 30736 MW; FFC4BB863DA0EDSF CRC64;

Query Match 95.8%; Score 115; DB 10; Length 281;
 Best Local Similarity 86.4%; Pred. No. 2,7e-11;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCCAGXKG 23
 Db 103 CPXXEKKNDRICTNCCAGXKG 124

RESULT 2
 Q40378 PRELIMINARY; PRT; 397 AA.
 ID Q40378

QY 2 CPXXEEKNDRICTNCCAGXKG 23
 ||| ||||| ||||| |||
Db 45 CPRSEKKNDRICTNCCAGTKG 66

QY	2	CPXXEEKNDRICTNCCAGXKG	23
Db	43	CPRSEKNDRICTNCCAGMKG	64

QY 2 CPXXEEKNDRICTNCCAGXKG 23
Db 45 CP RSE EKNDRICTNCCAGMKG 66

QY 2 CPXXEEKNDRICTNCCAGXKG 23
Db 45 CP RSE EKNDRICTNCCAGMKG 66

QY 2 CPXXEEKNDRICTNCCAGXKG 23
||| : ||||| |||
Db 84 CPRSEGNAENRICTNCCAGRKG 105

QY 2 CPXXEEKNDRICTNCCAGXKG 23
||| : ||||| |||
Db 84 CPRSEGNAENRICTNCCAGRKG 105

RESULT 6
 O82735 ID O82735 PRELIMINARY; PRT; 126 AA.
 AC O82735;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Potato (Solanum tuberosum) mRNA 2 for proteinase inhibitor II (Fragment).
 GN PIN2A.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sanchez-Serrano J., Schmidt R., Schell J., Willmitzer L.;
 RT "Nucleotide sequence of proteinase inhibitor II encoding cDNA of potato (Solanum tuberosum) and its mode of expression.";
 RL Mol. Gen. Genet. 203:15-20(1996).
 DR EMBL; X03779; CAA27409.1; -.
 DR HSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot_inhib.
 DR Pfam; PF02428; Prot_inhib_II; 2.
 FT NON_TER 1
 SQ SEQUENCE 126 AA; 13495 MW; 716A5A68DAFD57B2 CRC64;
 Query Match 65.0%; Score 78; DB 10; Length 126;
 Best Local Similarity 59.1%; Pred. No. 2e-05;
 Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
 Qy 2 CPXXEEKKNDRICTNCCAGXKG 23
 Db 18 CPRSEGSFQPICTNCCSGYKG 39

RESULT 7
 O82735 ID O82735 PRELIMINARY; PRT; 109 AA.
 AC O82735;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Proteinase inhibitor IIB (Fragment).
 GN PIN2B.
 OS Solanum americanum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=109975;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu Z.F., Qi W.Q., Ouyang X.Z., Yeung E., Chye M.L.;
 RT "Isolation of pepper RNAs differentially expressed during the hypersensitive response to tobacco mosaic virus and characterization of a proteinase inhibitor gene.";
 RL Plant Sci. 161:727-737(2001).
 DR EMBL; AF242734; AAF63518.1; -.
 DR HSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot_inhib.
 DR Pfam; PF02428; Prot_inhib_II; 3.
 FT NON_TER 1
 SQ SEQUENCE 109 AA; 11730 MW; 0CC6CCDASBA19F16 CRC64;
 Query Match 62.5%; Score 75; DB 10; Length 109;
 Best Local Similarity 54.5%; Pred. No. 5.6e-05;
 Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 Qy 2 CPXXEEKKNDRICTNCCAGXKG 23
 Db 3 CPRSEGSFQPICTNCCSGYKG 24

RESULT 8
 Q8W2K6 ID Q8W2K6 PRELIMINARY; PRT; 148 AA.
 AC Q8W2K6;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Proteinase inhibitor IIA.
 GN PIN2A.
 OS Solanum americanum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=109975;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu Z.F., Qi W.Q., Ouyang X.Z., Yeung E., Chye M.L.;
 RT "A proteinase inhibitor II of Solanum americanum is expressed in phloem.";
 RL Plant Mol. Biol. 47:727-738(2001).
 DR EMBL; AF174381; AAL36458.1; -.
 DR InterPro; IPR003465; Prot_inhib.
 DR Pfam; PF02428; Prot_inhib_II; 2.
 FT NON_TER 1
 SQ SEQUENCE 148 AA; 15323 MW; 053536378E104E75 CRC64;
 Query Match 60.8%; Score 73; DB 10; Length 148;
 Best Local Similarity 54.5%; Pred. No. 0.00016;
 Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
 Qy 2 CPXXEEKKNDRICTNCCAGXKG 23
 Db 42 CPRSEGSFQPICTNCCSGYKG 63

RESULT 9
 Q9M514 ID Q9M514 PRELIMINARY; PRT; 143 AA.
 AC Q9M514;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 15.7 kDa protein (Fragment).
 OS Capsicum annuum (Bell pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
 OX NCBI_TaxID=4072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shin R., Lee G.J., Park C.J., Kim T.Y., You J.S., Nam Y.W., Paek K.H.;
 RT "Isolation of pepper RNAs differentially expressed during the hypersensitive response to tobacco mosaic virus and characterization of a proteinase inhibitor gene.";
 RL Plant Sci. 161:727-737(2001).
 DR EMBL; AF242734; AAF63518.1; -.
 DR HSP; P01080; 4SGB.
 DR InterPro; IPR003465; Prot_inhib.
 DR Pfam; PF02428; Prot_inhib_II; 3.
 FT NON_TER 1
 SQ SEQUENCE 143 AA; 15658 MW; 401D0F1216CD2BE4 CRC64;
 Query Match 59.2%; Score 71; DB 10; Length 143;
 Best Local Similarity 60.9%; Pred. No. 0.00033;
 Matches 14; Conservative 1; Mismatches 6; Indels 2; Gaps 1;
 Qy 3 PXXEEKKNDRICTNCCAGXKG 23
 Db 3 PNASAKGNAENRICTNCCAGXKG 25

```

RESULT 10
P83241      PRELIMINARY;      PRT;      52 AA.
ID P83241
AC P83241
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DE Proteinase inhibitor PSI-1.2.
OS Capricornium annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
[1]
RN SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.
RP TISSUE=SEED;
RC MEDLINE=21490999; PubMed=11604534;
RA Antcheva N., Pintar A., Patchy A., Simonsites A., Bara E.,
RA Tchobanov B., Pongor S.;
RT "Proteins of circularly permuted sequence present within the same
RT organism: the major serine proteinase inhibitor from Capsicum annuum
RT seeds."
RL Protein Sci. 10:2280-2290(2001).
CC -1- FUNCTION: POTENT INHIBITOR OF TRYPSIN AND A WEAKER INHIBITOR OF
CC CHYMOTRYPSIN. IT DOES NOT INHIBIT ELASTASE AND SUBSTITISIN DY.
CC -1- MASS SPECTROMETRY: MW=5594.2; METHOD=ELECTROSPRAY.
CC -1- SIMILARITY: BELONGS TO THE POTATO TYPE II PROTEINASE INHIBITOR
CC FAMILY.
DR InterPro: IPR003465; Prot_inhib.
DR Pfam: PF02428; Prot_inhib_II; 1.
KM Serine protease inhibitor.
FT DISULFID 3 32
FT DISULFID 7 28
FT DISULFID 16 38
FT DISULFID 31 49
SQ SEQUENCE 52 AA; 5604 MW; 201823D8DA49CC61 CRC64;

Query Match 58.3%; Score 70; DB 10; Length 52;
Best Local Similarity 50.0%; Pred. No. 0.0002;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCCAGKKG 23
DB 16 CPSSGERIIRKVCNCCAGKKG 37

RESULT 11
O94192      PRELIMINARY;      PRT;      46 AA.
ID O94192
AC O94192
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DE Hypothetical 4.9 kDa protein (Fragment).
OS Atropa belladonna (Belladonna) (Deadly nightshade).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Atropa.
OX NCBI_TaxID=33113;
[1]
RN SEQUENCE FROM N.A.
RP Noar E.;
RA Theis (2001), Department of Plant Biotechnology,
RA Universite Libre de Bruxelles, Brussels, Belgium.
RL EMBL: AJ309380; CAC40750.1; -.
DR InterPro: IPR003465; Prot_inhib.
DR Pfam: PF02428; Prot_inhib_II; 1.
KW Hypothetical protein.

```

```

FT NON TER 1 1
SQ SEQUENCE 46 AA; 4881 MW; 0FFBF29C11BEEA00 CRC64;

Query Match 45.0%; Score 54; DB 10; Length 46;
Best Local Similarity 36.4%; Pred. No. 0.004;
Matches 8; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 CPXXEKKNDRICTNCCAGKKG 23
DB 10 CPSSGERIIRKVCNCCAGKKG 31

RESULT 12
O9M564      PRELIMINARY;      PRT;      417 AA.
ID O9M564
AC O9M564
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DE A6 gene product.
DE A6 OR EG:9D2.3 OR CG3771.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pezomacrida; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabrielsen A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glisler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Moberg A.,
RA Merklov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclid J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weissbrock G.M., Weissbrock J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster."
DR Science 287:2185-2195 (2000).
DR EMBL: AB003421; AAF45661.1; -.
DR FlyBase: FBgn0023130; a6.
SQ SEQUENCE 417 AA; 45479 MW; E9F7E35AE2D490DB CRC64;

Query Match 44.2%; Score 53; DB 5; Length 417;

```

Best Local Similarity 41.2%; Pred. No. 0.83;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 CPXXEKKNDICTNCC 18
DB 381 CPSGDGSKDLVCFRC 397

RESULT 13
Q9FG64
ID Q9FG64 PRELIMINARY; PRT; 361 AA.
AC Q9FG64
DT 01-WAR-2001 (TRENBLrel. 16, Created)
DT 01-WAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Similarity to Tail-like non-LTR retroelement protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026661; BAB09362.1; -;
DR InterPro; IPR001878; Znf CCHC.
DR SMART; SM00343; Znf C2HC; 1.
SQ SEQUENCE 361 AA; 41877 MW; BB84B28FOA398865 CRC64;

Query Match 43.3%; Score 52; DB 10; Length 361;
Best Local Similarity 61.5%; Pred. No. 1.1;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 BEKKNDICTNCC 18
DB 197 EYEKLQVCTNCC 209

RESULT 14
Q9AK62
ID Q9AK62 PRELIMINARY; PRT; 377 AA.
AC Q9AK62
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Putative alcohol dehydrogenase.
GN SC04055 OR 2SCD60.21C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Harris D.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96(1996).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomsen N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
FAMILY.
CC EMBL; AL583944; CAC32326.1; -;
DR HSSP; P14941; IYKF.
DR InterPro; IPR002328; ADH_zinc.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF0107; adh_zinc; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
DR Oxidoreductase; Zinc.
SQ SEQUENCE 377 AA; 40031 MW; 6EB9079B5854A28D CRC64;

Query Match 43.3%; Score 52; DB 16; Length 377;
Best Local Similarity 42.3%; Pred. No. 1.1;
Matches 11; Conservative 1; Mismatches 4; Indels 10; Gaps 1;

OY 8 KKDRI-----CTNCCAGXKG 23
DB 76 KKGDRVMPFVACGCTNCAAGYTG 101

RESULT 15
Q9LNG2
ID Q9LNG2 PRELIMINARY; PRT; 168 AA.
AC Q9LNG2
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE F21D18.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F21D18 from chromosome
I.";
RN [2]
RP Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

